

Db 318 TTTT... 377  
 Oy 417 GATATCTGGGGGTTTCTTATCTGATTTGATGACGTGATAGGTTAATCTT 476  
 Db 378 TTTT... 437  
 Oy 477 ACAGAGCTTATTTAGATATCAAGCTCTGTGATAAGATGCTCCAAAGCCGCTTTTAA 536  
 Db 438 NNN... 497  
 Oy 537 TGGATCTGGGGTAAAGAAATGATGTCGACGTGCTACTTACTATGCTGACAGAAATG 596  
 Db 498 KARK... 557  
 Oy 597 ACAGAAATTTATTTATCCGTTTACATTTGCTAATAGTATATAGAACTGAATAATAGC 656  
 Db 558 TGCT... 617  
 Oy 657 CTTAATATAGCTTTTGGAGGTTTGGAGCTTGTGATATTAATGAGGCAAGCTAGC 716  
 Db 618 AAGATTTTMAAATTTTGGGTTTGGAGATTAATTAAGACRGRTAGADAKTTAT 677  
 Oy 717 GCTTGGCTTGAACATCTTCAAAAAAGTTATTAATTTTATTTATTTAAATG 776  
 Db 678 GGAGCTTTT... 737  
 Oy 777 AGATATGCAATTTTGTATGTCATGTCATGTAATTTGAATTTGAGCTTTGTTA 836  
 Db 738 AGGATTTGDCAGTDTTGTGTTTGTGCTTTGAGTTTGAGRTVAGKATTTTGRTGTT 797  
 Oy 837 TTATTTATTTTATTT 852  
 Db 798 TGATTTATGCTTTT 813  
  
 RESULT 15  
 CDS03M11  
 LOCUS  
 DEFINITION  
 900 bp DNA linear GSS 17-MAY-2000  
 Tetradon nigroviridis genome survey sequence T7 end of clone  
 036N04 of library G from Tetradon nigroviridis, genomic survey  
 sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Tetradon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetradon.  
 1. (bases 1 to 900)  
 Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Benoit, A., Fizames, C., Wincker, P., Brothier, P., Quetier, F.,  
 Saurin, W. and Weissenbach, J.  
 Human gene number estimate provided by genome wide analysis using  
 Tetradon nigroviridis DNA sequence  
 2. (bases 1 to 900)  
 Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,  
 Bouneau, L., Billaud, A., Quetier, F., Saurin, W., Benoit, A. and  
 Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetradon nigroviridis  
 Unpublished  
 3. (bases 1 to 900)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000)  
 This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetradon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/tetradon.  
 Location/Qualifiers  
 FEATURES

source 1. 900  
 /organism="Tetradon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="036N04"  
 /note="Genoscope sequence ID : C08G0360S02LP1-end : T7"  
 BASE COUNT 226 a 154 c 115 g 340 t 65 others  
 ORIGIN  
 Query Match 7.7%; Score 66; DB 17; Length 900;  
 Best Local Similarity 42.6%; Pred. No. 0.035;  
 Matches 202; Conservative 21; Mismatches 251; Indels 0; Gaps 0;  
 Oy 3 TTTTTCATTGGATCCAGAAATTTGTAGAAATTTGACAAATTAAGACATTAATTAAGAA 62  
 Db 423 TTTCCTTTTTCCTCTTTTNTTTTAAATMACTCCCTATCCATCAATTAATTAATTAAT 482  
 Oy 63 TTGAATTTGCTAATTTTATGCTCAATCAACAGACCTGATTTGGAGGATTTCAAGTA 122  
 Db 483 AAAAAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 542  
 Oy 123 ATACTTGAATAAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 182  
 Db 543 TTTTATTTATACAMTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 602  
 Oy 183 TGTACTATTAATACAGTTTAAATTAAGAGCTTTTATGAATAATTAATTAATTTATGTT 242  
 Db 603 TTAHCAATTTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 662  
 Oy 243 TTGCTAGCCTTTTGTGTTTCTTCTAGCTTCTCAAGCTGATGATTTCAAAAAGGCTTT 302  
 Db 663 TAAATAAATTTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTT 722  
 Oy 303 AATTTGGAGGAGGGAATAAATTTTACCTATGTAAGACTAGTAAGAAAGATCTATTTG 362  
 Db 723 AATTAATTTACATTTATTAATTAATTTTATTAATTAATTTTATTAATTTATTTAT 782  
 Oy 363 CCATTTTATTTGAACCTTTTATAGGTTTGAATAGCTTTTGTCTCAAGAGATATTT 422  
 Db 783 AAAACTTTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 842  
 Oy 423 CTGGGGGTTTCTTATTTCTGATTTGATGACGTGATAGGTTTATTTATTTATTT 476  
 Db 843 CTTT... 896

Search completed: March 27, 2003, 03:54:13  
 Job time : 1217.71 secs

```

RESULT 13
CONSOL6E2/c      CINSO16E2       1204 bp   DNA        linear    GSS 26-JUL-1999
LOCUS            Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION       BCN15A12 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION        ALI06628
VERSION          ALI06628.1  GI:5622852
KEYWORDS         GSS.
SOURCE           Drosophila melanogaster.
ORGANISM         Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epiphytoidea; Drosophilidae; Drosophila.
REFERENCE        1 (bases 1 to 1204)
AUTHORS          Direct Submission
TITLE            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr JOURNAL         Web : www.genoscope.cns.fr)
FEATURES         Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES         Location/Qualifiers
Source           1..1204
                 /organism="Drosophila melanogaster"
                 /db_xref="taxon:7227"
                 /clone="BACN15A12"
                 /clone_lib="DrosBAC"
                 /plasmid_id="pBelobAC11"
                 /note="end : 77"
BASE COUNT       298 a 172 c 106 g 316 t 312 others
ORIGIN
Query Match     7.7% Score 66.6 DB 17 Length 1204; Best Local Similarity 21.3%; Pred. No. 0.025; Matches 107; Conservative 181; Mismatches 215; Indels 0; Gaps 0;
QY 22 ATTGTAGAATTTTCGACAATAATTAAGACATTATTAAGAATTGAATTGCTATTTTAT 81
Db 1200 ADARAAAMWRAKAKMKWKMTAAAWMAKAKRKAHAIAAAAAAAAAAAAAADNA 1141
YY 82 GGTCAATCAAGAACCTCTATTGGAGACGCAATTTCAAATAACTTGAAAAGTTAA 141
Db 1140 AAKKKAAAAAAAKAADDAADAAAAAKAKAKAKAAAAAAADAAMKKKAAAAAAKWA 1081
QY 142 ATTTAAATAGTTTTAAACCCTTTTAAATTTCAATTATAGTAGCATATACCAGT 201
Db 1080 AAKMWTATRKDDDDDKRAAADAKKKKKKAPADRADKDCKAAADWDADDAAKKK 1021
QY 202 TTAAATAAAGAGCTTTTATGAATAATTTTAAATTTTATTTGTTTGCCTTTTGT 261
Db 1020 KAKKKDKKDATADDMAAKAAKKDKRAMKKKXKKAKAKAKKKDKDKDKKKKK 961
QY 262 TTTTCAGCTTTGGCTCAAGCTGATGTTCTTAAGAAGCGCTTTTAATTTGGAGC 321
Db 960 KKDKKAKKKKRAAADAAAAAAAKKKDKKDAKAAKAKKKKDKAARKKKKDDAMA 901
QY 322 AACCTTTGCTTATCAACTAGTAGAAGATCCATTTGTGCCATTTTATGAACCTTT 381
Db 900 ADAADKAAAKKRAAKAKKKKKADADKAAKADAKKKKKRAAKAADKDAKMDKKKAG 841
QY 382 TTTTAAAGGCTTGAATAGTCTTTTTCCTCAAGAGATATCTTGAGGCTTTTCATTC 441
Db 840 AAAAARAAGGCKKGAGGAGGAAAGATTTWAKTDJTKTKAGDAKTITTKTTTWA 781
QY 442 TTGGATTGATGCACTGCTGTATAGGCTTAATACCTTACAGGAGCTTATTTAGATCAAG 501

```

Db	780	TKAGKMKKTRRAAKAKDARKGDAATTAKAARAGAAATATGGGAARAGGAAGGAKDAGARAK	721
OY	502	CTCTGATAGAAATGCTCCAAA	524
Db	720	GKAAAKAGATKGAAAATACMWA	698
RESULT 14			
CNS00YXQ			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
OY	57	AAAGATGAATGCTATTTTATGCAATCAAGAAGCTATTTGGAAGCGAATTT	116
Db	28	VWMAWMTWMSATACOTTTTTTTTCKTWTNTTTTTTTTTTNNNNNTNNNTT	87
OY	117	CAAGTATACCTTTGAAAAAGTTAAATTTAAATGTTTAAAAACCTTTTAAATTCA	176
Db	88	NTNNTCNTNTTNRCCNNNNNTNNNNAAAAAGWAACTAAAAAAATTTTNTNTNTN	147
OY	177	TTAATAGTACTAATATACAGTTTAAATAAGAGTTTAAATGAATAAATTTTAAT	236
Db	148	TTTTTNNNTNTNTANANNAANNNNTNTTTTTTTTTTNTNTTTTTTTTTTTTTT	207
OY	237	GTTGTTTTGACGCTTTTGTGTTTCTAGCTTTCCTCAAGCTGATGCTTAAAGC	296
Db	208	TT	267
OY	297	GCTTTAAATTTGGAGCGGAGAAAACTTTTACGTTATGAAACTAGTAAGAAAGATCT	356
Db	268	TT	317
OY	357	ATTGCGCAATTTTATTAACCTTTTTTTTACGGTTTGAATAGGCTCTTGCTCAAG	416

filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
location/Qualifiers

source

source

```

source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACROBE08"
/clone_1b="RPC1-98"
/note="end : TERT3"
553 a 143 c 100 g 134 t 171 others
BASE COUNT
ORIGIN

```

Query Match	7.7%	Score 66.8;	DB 17;	Length 1101;
Best Local Similarity	38.3%	Pred. No. 0.024;		
Matches 184;	Conservative 46;	Mismatches 250;	Indels 0;	Gaps 0;

```

OY      36 AAAAAAGAAATGCAATTCGAATTATGGCTCAACATCAGAAGCCATCATATGGAGAACCAATTTT
Db      918 AAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTW
OY      116 TCAAGTAATACTTTGAAAAAAGTAATTAATTAATAGTTTTAAAAACCTTTTAAAATTTTC
Db      858 TTTTTTTTTTTTTTTTTTTNNATATWTWWAAAAAAAYTTTTTTNNAATTTWTTTTTTTTTTT
OY      176 ATTAATAGTACTATAATAACCACTTTTAATAAAGAGTTTTATGATTAATTAATTTTAAAT
Db      798 TTTTTTTTTTTTMMMAAAAWTTTTTTTTTARTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
OY      236 TGTTGTTTGCTACCGTTTGTGTTTTTTTCTACGTTTGCACAAGCTGATTTCTAAAG
Db      738 TTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTKGGRT
OY      296 CGCTTTAAATTTGGAGGAGGAGAAAACTTTTGCTATGGAACATGTAAGAAGATCC
Db      678 KRGKGGGCKKKKGCGGCTKKGCGGCTTTTTTTTAAATTTTTTTTGTGGGGGCGKT
OY      356 TATGTGCCATTTTATTTGAACCTTTTTTTTTAGGCTTTGAATAGTCTTTTCTCAAG
Db      618 TTTTTGAATTTTTTTTTTTTCKTTLTKKKKGCGGGGCKKTTTTTTTTTKGTTTTGG
OY      416 AGATATTTCTGGGGGTTTCTTATTTCTTGATTTGATGCAGTTGCTATAGGCTTAATACT
Db      558 GKTATTTTTRKTGATTTTTGGTGGGCGKCKKTKTKTTTTTTTTTTTTTGGGTTTTGCG
OY      476 TACAGACCCTATTAGATATCAAGACGCTGTGATAAAGATGCTCAAACCGCGTTTAA
Db      438 TTTTNTTCTTTTGGGGGARVRARRTTTNGGRRPAAARVMMMMMTTHSCAATTNAAA

```

LOCUS	CNS006VL/12	1101 bp	DNA	linear	GSS_04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence 17 end of BAC: BAC239623 of RP11-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL069706				
VERSION	AL069706.1 GI:4949849				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information				

please see [http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library](http://www.fruitfly.org/The%20BDGP%20Drosophila%20melanogaster%20BAC%20library) was prepared by Kazuhiro Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

Source

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RCR1-98"
/note="end : 77"

```

Query Match	7.7%	Score	66.6	DB	17	Length	1101
Best Local Similarity	30.9%	Pred.	No. 0.026				
Matches 205; Conservative	133; Mismatches	319;				Indels	6; Gaps
							2;

QY	22	ATTGTGCAATTTTCGCAAAATTAAGACATATATTAAGAAAGTAATGCTATATTTAT	81
Db	1036	ATTTTATTTTATTTTATTTTATTTTAAATATATTAATTAATTTTATTTATTTATTTATTA	103
QY	82	GCTCAATCAAGAAGCTCTATTGGGAAGCGAATTTCAAGTAATACCTTTGAAAAAGTTAA	141
Db	1036	TATTTATTTTAAATATATATATTTAATTAATATATTTATTTATTTATTTATTTATTTAT	977
QY	142	ATTTAATAGTTTAAATTAACCTTTTAAATTCATTAATATGCTACTATAATACAGTT	201
Db	976	AAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTTATTTT	917
QY	202	TTAATTAAGAGCTTTTATGAATAATTTTAAATTTGTTTGGTCTAGCCCTTTGCTTT	261
Db	916	TTTATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	862
QY	262	TTTCTAGCTTTCCTCAACCTGATCTCTAAACCCGCTTTAATTTGGAGCGGAGAAA	321
Db	861	ATTAATATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	802
QY	322	AACCTTTAGCTATGAACTAGTAAGAAAGATCCATTTGCGCACTTTTATTTGAACCTTT	381
Db	801	TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	742
QY	382	TTTATGAGCTTGGAAATGAGCTTCTTTGCTCAAGAGATATTCTTGGGGTTTCTTATTC	441
Db	741	TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	603
QY	442	TGGAATTTGATGCGACCTGCTATAGGCTTAATACCTTACGAGAGCTTATTTAGATCAAG	501
Db	682	AATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	623
QY	502	CTCTGATPAGAGTCTCCAAAAGCCGCTTTAAGTGACCTTGGGGTGAAGGAATGATGT	561
Db	622	AAAAATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	563
QY	562	TGGCAGGCGACCTGATCTATAGGCTGACACAGTTGACGAAATTAATTTACCTTACAT	621
Db	562	TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	503
QY	622	TTCCTAATAGTATATAGAAACTGAAAAATAGCCTTAATATAGCTTTTGGAGGGTTTG	681
Db	502	TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	443
QY	682	AGC 684	
Db	442	GGG 440	

[illegible]

RESULT 10	CNS0005X	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
	CNS0005X	1101 bp	DNA	linear	GSS 03-JUN-1999		
	Drosophila melanogaster	genome survey sequence	TEF3 end of BAC #				
	BA0R02L06 of Rpcl-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.						
	AL050813						
	AL050813..1	GI:4930826					
	GSS.						
	Drosophila melanogaster.						
	Drosophila melanogaster						
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.						
	1 (bases 1 to 1101)						
	Genoscope.						
	Direct Submission						
	Submitted (02-JUN-1999)	Genoscope - Centre National de Sequencage	:				

**COMMENT**

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oseegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```
FEATURES      Location/Qualifiers
source        1. .1101
```

```

BASE COUNT      77 a      768 t      145 others
ORIGIN          26 c      85 g

```

Query Match	7.8%;	Score 67.6;	DB 17;	Length 1101;
-------------	-------	-------------	--------	--------------

Matches 287; Conservative 40; Mismatches 385; Indels 2; Gaps 2;

138 TTAATTAAATAGTTTAAACCTTTTAAATTTCATTAAATGTTACTATAATACC 19

DB LU / I.I.I.I.I.KA.I.I.I.I.I.NNNNNNNNNI.TTTT.TTTT.ATTTTTTTTT 166

[illegible][illegible]

```
Db      227 TTTTTTTT|TTTTTTT|TTTTTTT|TTTTTTT|TTTTTTT|TTTTTTT|TTTTTTT|TTTTTTT|
          ||||| |   ||| |   ||| |   ||||| ||||| 
          286
```

318 GAAAACTTTAGCTTATGAACTAGTAAGAAAGATCCTATTGTGCCATTTTATTGAC 377

[illegible]

3/8 CTTCTTCTAGGGTCTGGAAATAGGTTCTTCTCTCAAGGAGATATCTTGGGGTCTCTT 43/

[illegible][illegible]

**THE UNIVERSITY OF CHICAGO**

[illegible]

017 558 AAAAAAAAAAAAAAAACCAAAAACCAAAAAAATTTTAAAAAACCTTTT 617

Dh 527 GATTCTTTTTCWRTTGTRGTTGTAGTTTTGGGTTTTRGTRGTDWRGTTTTATTTGRTTGG 586

618 ACATTGCTAATAGTTATAATAGCAACTGAAAAATAGCCTTAATATAGCTTTGGAGG 677

Db 587 TTTTWTATATATAATTKRTTGTGARTTGGATRTTTTGGATTTT 646

678 TTGAGCCTAGTTTGATATTAATATGGGCCAAGCTAGCGCTTGGGTTGA  
737 TGA

Db 647 KTTGKTAGTRARTT-ATAGTKWTTTATTKTKGWTGTTRTTGGGTTTTTTTGGTTTTT 705

738 TTCAAAAAAGCTTATTATTATTATTATTAAATGAGTGATAGCAATTTGTATTG 797

[illegible][illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

RESULT 11

LOCUS	1101 bp	DNA	linear	GSS 03-JUN-2003
CNS003B0				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of AAC			

BACR08E08 of RPCL-98 library from *Drosophila melanogaster* (fruit fly); genomic survey sequence.

ACCESSION	AL064078
VERSION	AL064078.1
	GI:4941834

SOURCE *Drosophila melanogaster*.

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Ephydroidae; Drosophilidae; Drosophila.  
1 (base 1 + 1101)

AUTHORS	Genoscope.
TITLE	Direct Submission

JOURNAL  
Submitted (02-JUN-1999) Genoscope - Centre National de séquençage  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cnrs.fr](mailto:secref@genoscope.cnrs.fr))

web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))  
Determination of this BAC-end sequence was carried out as part

The BDGP is constructing a physical map of the *Drosophila*

please see <http://www.fruitley.org> The BDGP Drosophila

Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo

NY. The library is named RPCI-98 and was constructed by partial digestion of *Drosophila* DNA provided by the RING from the

isogenic strain y2; cn dw sp, the same strain used for the Bbci  
P1 and EST libraries. A more detailed description of the libra

WOMEN  
TO  
TECHNICAL  
INDUSTRIAL  
AND  
CONSUMER, THE  
CULTURE  
INDUSTRY, OR



Db	1084	TTTAAATAATTTTTTTTTTAATTATTTTTTTATTTTTTTTTTTTTTTTTTANNTTTT	1138
RESULT 8	CNS015WU/c	1203 bp DNA linear GSS 26-JUL-1999	
LOCUS	CNS015WU	Drosophila melanogaster genome survey sequence SP6 end of BAC	
DEFINITION	BACN15E10 of DrosBAC library from Drosophila melanogaster (fruit fly); genomic survey sequence.		
ACCESSION	AJ106008		
VERSION	ALI06008.1	GI:5619558	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1203)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosBAC11.		
FEATURES	Location/Qualifiers		
source	1..1203		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone_lib="BACN15E10"		
	/clone_lib="DrosBAC"		
	/plasmid="pBelosBAC11"		
	/note="end : SP6"		
BASE COUNT	288 a 103 c 235 g 253 t 324 others		
ORIGIN			
	Query Match 7.9% Score 67.8; DB 17; Length 1203;		
	Best Local Similarity 40.2%; Pred. No. 0.016; Indels 0;		
	Matches 96; Conservative 48; Mismatches 95; Gaps 0;		
	21 AATTGTGAGAAATTTTCGACAATAAAGAATTTTAAAAAGAAATGGAAATTGCTAAATTTTA 80		
	:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::		
	1132 WATWTATATATWTATTTTWTATTTATTTTATTTTATTTTATTTTAAAAMAWMAATTTTWWAMA 1073		
	OY 81 TGGCGAATCAGAAGAGCTCTATTTGGGAAGCGAATTTCAAGTAATACTTTGAAAAAAGTTA 140		
	:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::		
	Db 1072 TWATMAATWAAAAAAMWTTATTAAMWMAAMATAAAMAWMAAMTWMTATATTAATAAATAAT 1013		
	OY 141 AATTTAAATAGTTTAAAAAACCTTTTAAAAATTCATTAATTTGTTACTATATACCAGT 200		
	:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::		
	Db 1012 AATWTAAMAWMTWTWTAAMAAATTTTAAAAAAMAAAMAAWMAATATATATATAAT 953		
	OY 201 TTTTAATTAAGAGCTTTTATGATATAAATTTTAAATGTTGTTTCTAGCCCTTTGTGT 259		
	:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::		
	Db 952 SATWTATWMAATTTTWTTTTATTAATTTTTCCTATATWATCAMCCCCCYTTTMTWT 894		
RESULT 9	CNS005F3/c	960 bp DNA linear GSS 03-JUN-1999	
LOCUS	CNS005F3	Drosophila melanogaster genome survey sequence TEK3 end of BAC #	
DEFINITION	BACR1P07 of RPII-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.		
ACCESSION	AL059925	GI:4943047	
VERSION	AL059925.1	GI:4943047	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
Drosophila melanogaster	Eukaryota: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	1 (bases 1 to 960)	Genoscope.	Direct Submission
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr - Web : www.genoscope.cns.fr)			
	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org/The%20BDGP%20Drosophila%20melanogaster%20BAC%20library">http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library</a> was prepared by Kazuhiro Oosawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .			
FEATURES	source	1..960		
	/organism="Drosophila melanogaster"			
	/db_xref="taxon:7227"			
	/clone="BACR11P07"			
	/clone_11b="RPc1-98"			
	/note="end : TET3"			
BASE COUNT	570 a	112 c	37 g	71 t 170 others
ORIGIN				
Query Match	7.8%.	Score 67.6:	DB 17;	Length 960;
Best Local Similarity	35.4%.	Pred. No. 0.018;		
Matches 253;	Conservative 88;	Mismatches 369;	Indels 5;	Gaps 1;
QY	143	TTTAAATAGTTTAAACACCTTTTAAATTCATTAAATGTACTATAATACAGTTT	202	
DB	853	TTGTAAKARKTDTTTRCANLTKTKTGRGTATTTTCTGTGTTTAAATGTTTAAAG	794	
QY	203	TAATTAAGAGCTTTTATGAATAAATTTTAAATGTGTGTTGGACGCTTTTGCTTT	262	
DB	793	GTTTGARAGTAKARTAAATGAATTTGGTGGAAGAGCTGKTKTRAAATTTGDTTGKTTG	734	
QY	263	TTTCAGCTTTTGCTCAAGCGATGATCTTCAAAAGCCCTTTTAAATTTGGAGCGGAGAAA	322	
DB	793	DWTWTTTTT-----TTGTGTTGGTGGARWAAATKTRRAADPTAATAKTGRMAT	679	
QY	323	ACTTTAGCTTATGAACACTAGTACAAGATCCATTTGTCATTTTATTAACGCTTTT	382	
DB	678	TTTTTTATTTRAMTGTGTGTTTAAATTAATGTTTDTTDTTATTTTATTTTCTTGT	619	
QY	383	TTTACGTTTGGAAATAGGCTTTTGGTCAAGACAGATATCTTTGGGGGTTTCTTATTC	442	
DB	618	KAGAGAGDGGARAKRKTITTKTTTGGGGAATTTKGTGTTTRATMGARPTTKTKIST	559	
QY	443	TGGATTGATGACAGTTTGGTATAGGGTTTATACCTTACAGAGCTTATTTAGATATCAAGC	502	
DB	558	TGTTTTTKTKTGTTTTTTGTTTGGTKGTGGAARKTKGTGTGTGRTGTGTGRAD	499	
QY	503	TCTTATTAAGAATGCTCCAAAAGCCGCTTTTAAAGTGACCTTGGGGTAAGGATATGTT	562	
DB	498	AAAATTTTGAADGKGCTTTGTTTCTTTTCTGCTGMAATGKTKGRKTGKTKGRDTC	439	
QY	563	GCCAGCTGACAGTTATAGGCTGTGACACAATGACAGAAATTAATATCCGTTACATT	622	
DB	438	TTTTTTTDTGKGAATTTDTTKTKADWTTTTTGKKAAGRTTTTTTTTTTTTTNTTT	379	
QY	623	TGCTAATAGTATTAAGGAAACGAAAAATAGCCTTAATATAGCTTTTGGAGGCTTGA	682	



```

Db 1034 TWDTWMDMMWMAATAKIDTAMTWRTAMRADWAGRDGAGKRDADAATDGDAGRRGG 975
Oy 296 CGCTTTATTTGGAGCGGAGAAAACCTTTAGCTATGAACTGAAGAATCC 355
Db 974 RKRKRKRKGGDDKGGKKKAAKATKMDMDKMDKMDKMDKMDKMDKMDKMDKMD 915
Oy 356 TATGTGCAATTTTATGAACTTTTGAAGGTTGGAATAGCTTTTGTCTCAAG 415
Db 914 GKDGDGKGKADDDTDTGKDDKDKMDMDKAGTGDATWMAAATDMMWMAADAD 855
Oy 416 AGATTTCTTGGGGTTTCTTATCTTGATTTGATGACAGTGTATAGGTTATACT 475
Db 854 WTWDAADWMAADWMAADWMAAMKMDAMWARTADRDWGRGARGKARRRRRAD 795
Oy 476 TACAGACTTATAGATATCAAGCTCTGATAGAATGCTCAAAAGCCGCTTTTA 535
Db 794 KRDAADDDADATTTTTRTDMDKWKTDTRMAADTTMDRDDDD-DRDRACTAG 736
Oy 536 GTGACTTGGGTAAGGAATGATGTGGCAGTGCACTTACTATGCTGACAAAGATT 595
Db 725 RKRRTWRKRRKRRRTDMDADADTDARDRRRRGGDAGKGGKRRRRRRATWD 676
Oy 596 GACAGAATTTATTTCCGCTTACATTTGCTATATGTAATAGGAACGAAATAG 655
Db 675 RTDAMWMAAAMWTTTDTDDMDKRRRRKRRRRRTTARAAMDMWMAKMDK 616
Oy 656 CCTATATAGCTTTGAGGTTGAGCCTAGTTGATATATGAGGCGCAAGCTAG 715
Db 615 TRADWMDWMAADTTDAKADBDMAKARARARARARARARARARARARARAT 556
Oy 716 CGCTCTGGTGTGAATCTTTCAAAAAGTTATTAATTTTATTTATTAAT 775
Db 555 AARAAWMAWMAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 496
Oy 776 GAGTATAGCAATTTGATGTGATGCTGCTATGTAATGAAATAGAGTTTGT 835
Db 495 AAAAAAAMWMAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 436
Oy 836 ATTATTTATTTTATTTCT 855
Db 435 WTTWMAATTTATTTTWTWT 416

RESULT 5
CNS04P4P/c 1184 bp DNA linear GSS 24-MAY-2000
LOCUS 125F20 of library G from Tetraodon nigroviridis, genomic survey
DEFINITION
ACCESSION AL300850
VERSION AL300850.1 GI:8177869
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1184)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizesen,C., Wlucker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
HUMAN gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 1184)
JOURNAL Roest-Crollius,H., Jallion,O., Dasilva,C., Fizesen,C., Fisher,C.,
REFERENCE Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
AUTHORS Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished

```

```

REFERENCE 3 (bases 1 to 1184)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1184
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="125F20"
/location="G"
/note="Genoscope sequence ID : COBG125DC10LP1-end : T7"
BASE COUNT 423 a 218 c 225 g 253 t 65 others
ORIGIN
Query Match 8.1%; Score 69.6; DB 17; Length 1184;
Best Local Similarity 44.4%; Pred. No. 0.0076;
Matches 120; Conservative 31; Mismatches 119; Indels 0; Gaps 0;

Oy 3 TTTTCATTTGATCCAGATTTGTAGATTTTGCACAATTAAGACATTTAAAGAA 62
Db 1081 TTTNTTATATATATATATATATATATATATATATATATATATATATAT 1022
Oy 63 TTGAATTCCTAATTTTATGTCATCAATCAAGAGCTATGGAAGCAATTCACATA 122
Db 1021 TTTTAAWTTTATTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTT 962
Oy 123 ATACTTAAAAAGTTAAATTTAAATAGTTTAAATTTTAAATTTTAAATTTTAA 182
Db 961 ATWTTTATATATATATATATATATATATATATATATATATATATATAT 902
Oy 183 TGTCTATATATACAGATTTTAAATAGCTTTTAAATATATATATATATAT 242
Db 901 TATTTTATATTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTT 842
Oy 243 TTGCTAGCCTTTGCTTTTCTAGCTTT 272
Db 841 TWTTTTAAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 812

RESULT 6
CNS00DKY 928 bp DNA linear GSS 04-JUN-1999
LOCUS BACR27A24 of Rpci-98 library from Drosophila melanogaster (fruit
DEFINITION fly), genomic survey sequence.
ACCESSION AL071865
VERSION AL071865.1 GI:4948170
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pelegrina;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 928)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequence;
REFERENCE BP 191 91006 EVR cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
AUTHORS - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org/The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oseegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named Rpci-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the

```

[illegible]

Dd 727 ATATATTCTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTKKTTAKRADDDDDKARRDCKGDMND 786

Oy 422 TCTTGCGGGGTTCCTTAATCCCTGAATTGCAGCAGTGGTAGGCTTAATCACTACAG 481  
:: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: :  
Db 787 ADKWKKAGHRDRDDTRKMWMKDRAKKKKKHKTAMWMMADDTGMTAAWTATAKTACT 846

Oy 482 ACCTAATTTAGCATATCAAAGCTCTTGATAGAANCGTCCAANAAGGCCGTTTAACTGAC 541  
:  
Db 847 GAAKATKBCKRRGGGRGDGFRGARFKRKARKKRGRRKRKKRKKRKKKGDSGTGAG 906

Oy 542 TTGGGGTAAGGAATGATGTTGGCAGGTCATCTACTATGCTGTGACAAGATTACAGA 601  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 907 TDKADKKKABDDSGMKRMKMDSKMTKTMGTDTDITTWKGNRRGRGAGRGRGRKRG 966

Oy 602 AATTAATTTATCCGTTTACCATTTCCTAATATGTAATAAGAAAAGCTGAATAAG 655  
- - - - - :  
Db 967 KAGKAGAGGWMAWADAEMTWTMVKDKWKMTDMCWGMGSDGRKRGKAGDG 1020

RESULT 4  
CNS0039G/c            1101 bp       DNA          linear      GSS 03-JUN-1999

LOCUS  
DEFINITION     Drosophila melanogaster genome survey sequence TEI3 end of BAC #  
BACR08k10 of RPci-98 library from drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION     AL063921  
VERSION       AL063921.1 GI:4941778  
KEYWORDS  
SOURCE        GSS.  
ORGANISM      Drosophila melanogaster.  
  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Epiphytroidae; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequeifgenoscope.cns.fr)

- Web : www.genoscope.cns.fr )  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mamosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the Isogenic strain Y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES  
source Location/Organism  
1..1101 /organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACR08K10"  
/clone\_lib="RPci-98"  
/note="end : TEI3"

BASE COUNT     201 a            64 c        131 g            202 t        503 others

ORIGIN

Query Match            8.8%; Score 75.6; DB 17; Length 1101;  
Best Local Similarity 17.6%; Pred.No.0.0007;  
Matches 120; Conservative 288; Mismatches 271; Indels 1; Gaps 1;

Oy 176 ATTATATGTCACATTAATCACAGTTTAATTAAGAGGTTTATGATTAATTTTAAAT 235  
::||:||:::||||::: ||::: ||::::||::: ||::: ||::: ||:  
Db 1094 DDYWDNRDJRDKNDBMYTKWWTYMKDAADRBAAGDADRBAWGDSGYTWTATYWMMWWWWA 1035

Oy 236 TGTTGTTTTGCTAGCCCTTTTGTTGTGTTTTTTCAGCTTTGCTCAAGCTGATGATTTCAAAG 295

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

Location/Qualifiers

1. 1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACR48p19"  
/clone\_lib="RPCT-98"  
/note="end : TET3"

BASE COUNT 469 a 6 c 69 g 151 t 406 others

ORIGIN

## Query Match

Best Local Similarity 24.4%; Score 84.4; DB 17; Length 1101;  
Matches 176; Conservative 216; Mismatches 328; Indels 2; Gaps 1;

```

38 ACAATTAAGACATTTTAAAGAAATGTAATTTTATGTCATCAATCAGAAGC 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 TCTATTGGAGCGAATTCAGTAATCTTGAAGAAAGCTTAATTTAAATAGTTAA 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 AAACCTTTTAAATTCATATATGTAATCTTAATCAGATTATTAAGAGTTT 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 TTTTWTWTTTDAADKMAAWDTTWTWTTWTTWTTWTTWTTWTTWTTWTTW 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 TATGATTAATTTTATGTTTGTGTCAGCTTTTGTGTTTGTGTTTGTGTC 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
562 WTKAKTAAKTAAMAAMTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278 AGCTGATGATCTAAAGCGCTTTAATTTGGAGCGGAGAAACCTTTTAACT 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
622 TTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 AACTAGTAAGAATCCATTTGTCATTTTATGAACCTTTTAAAGGTTTGAAT 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
682 AAAMWTTTDTMAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 741
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
398 AGCTGTTTGTCAAGAGATATCTGGGGGTTTCTTATCTTGATTTGATGAG 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
742 KKRKRGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 TGGTATAGGTTAATCTTACAGAGCTTTTAAATCAAAAGCTCTTGATAGA 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 TTTTGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 TCAGAAAGCGCTTTAAGGAGATCTGGGTAAGGAATGATCTGACAGTGTAC 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
862 KKKKKKKG--GKKKDAAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TATGCTGTGACAGATGACAGAAATTTATTCGTTTACATTTGCTAATAGT 637
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
920 KKKKGGGGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
638 TAGGAACGAAATAAGCTTAATAGCTTTTGAAGGTTTGAGCTTATGATAT 697
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
980 WTDAAATKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1039
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 TAATATGGGCAAGTACGCTTGTGGTTTGAATCTTTCAAAAAAGTTATTA 757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1040 KKTGKDKGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 1099
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
758 TT 759
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1100 KK 1101

```

RESULT 2  
CNS035N7 576 bp DNA linear GSS 15-MAY-2000  
LOCUS Tetradon nigroviridis genome survey sequence pUC-ori end of clone  
DEFINITION 214A06 of library G from Tetradon nigroviridis, genomic survey  
sequence.

ACCESSION AL228940  
VERSION AL228940.1 GI:7887933  
KEYWORDS GSS: genome survey sequence.  
SOURCE Tetradon nigroviridis.  
ORGANISM Tetradon nigroviridis.

Eukariota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.

1 (bases 1 to 576)  
Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Fizames, C., Munkner, P., Brotlier, P., Quetler, F.,  
Saurin, W. and Weissenbach, J.  
Human gene number estimate provided by genome wide analysis using  
Tetradon nigroviridis DNA sequence

2 (bases 1 to 576)  
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,  
Bouneau, L., Billault, A., Quetler, F., Saurin, W., Bernot, A. and  
Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetradon nigroviridis

Unpublished  
3 (bases 1 to 576)  
Direct Submission

Submitted (12-APR-2000)

This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetradon nigroviridis  
genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetradon>.

## FEATURES

source

Location/Qualifiers

1. 576  
/organism="Tetradon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="214A06"  
/clone\_lib="G"  
/note="Genoscope sequence ID : C0AG214A03SP1-end :  
pUC-ori"

BASE COUNT 62 a 5 c 18 g 444 t 47 others

ORIGIN

## Query Match

Best Local Similarity 44.3%; Score 78; DB 17; Length 576;  
Matches 162; Conservative 31; Mismatches 173; Indels 0; Gaps 0;

```

112 AATTCAAGTAATACCTTGAAGAAAGTAATTAATAGTTTAAACCTTTTAA 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 AATGAATAATTAAGATTTAATAAANNAATTTTAAATTTTAAATTTTAA 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 TTTCAATTAATGCTACTAATATACAGTTTAAAGAGGTTTATGATTAAT 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 TTTTWTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 TAATGTGTTTGTGACGCTTTTGTGTTTTCAGCTTGTGCTCAAGCTGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 TAAWMTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 AAAGCGCTTAATTTGGAGCGGGAAGAACTTTTAAAGTATGAACTAGTA 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
228 TTTTWTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 ATCCATATGTCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
288 GGGTTTWTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTT 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 AAGGAGATATCTTGGGGGTTTCTTATTTCTTGATTTGATGAGTTTAA 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1209.71 Seconds  
(without alignments)  
11540.380 Million cell updates/sec

Title: US-09-508-487-20

Perfect score: 862  
Sequence: 1 gattttcatgatccacg.....atatattctctcgtctaa 862

Scoring table: IDENTITY\_NIC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.4	9.8	1101	17	CNS00L72
2	78	9.0	576	17	CNS0035N7
3	77.8	9.0	1101	17	CNS0100X
4	75.6	8.8	1101	17	CNS0039G
5	69.6	8.1	1184	17	CNS04P4P
6	69.2	8.0	928	17	CNS00DKY

7	68.2	7.9	1167	17	CNS07360	AL427102 clone BAO
8	67.8	7.9	1203	17	CNS015WU	AL106008 Drosophila
9	67.6	7.8	960	17	CNS005F3	AL059925 Drosophila
10	67.6	7.8	1101	17	CNS000SX	AL050813 Drosophila
11	66.8	7.7	1101	17	CNS003BL	AL064078 Drosophila
12	66.6	7.7	1101	17	CNS00EVL	AL069706 Drosophila
13	66.6	7.7	1204	17	CNS016E2	AL106628 Drosophila
14	66.4	7.7	1204	17	CNS00YXG	AL096968 Drosophila
15	66	7.7	900	17	CNS03M11	AL250174 Tetradon
16	65.6	7.6	524	17	CNS01090	AL167541 Tetradon
17	65.4	7.6	570	12	BG680408	BG680408 1024070D0
18	65.2	7.6	1101	17	CNS00E07	AL069440 Drosophila
19	64.8	7.5	996	17	CNS00E0H	AL071063 Drosophila
20	64.6	7.5	862	17	CNS029ER	AL187164 Tetradon
21	64.6	7.5	1225	17	CNS0161D	AL106171 Drosophila
22	64.6	7.5	1303	14	BM924877	BM924877 AGENCOURT
23	64.4	7.5	893	17	CNS013XE	AL103436 Drosophila
24	64.4	7.5	915	17	CNS008BP	AL051471 Drosophila
25	64.2	7.4	1094	17	CNS012FZ	AL101513 Drosophila
26	64	7.4	1135	17	CNS033GO	AL226115 Tetradon
27	64	7.4	1200	17	CNS016CO	AL106578 Drosophila
28	63.8	7.4	669	17	CNS010M7	AL099143 Drosophila
29	63.6	7.4	916	17	B09084	B09084 F6L16-T7 IG
30	63.2	7.3	865	17	BH180441	BH180441 016_H-20-
31	63.2	7.3	865	17	CNS07MAN	AL617393 T3 end of
32	63.2	7.3	1101	17	CNS00L72	AL078714 Drosophila
33	63.2	7.3	1201	17	CNS0167M	AL106336 Drosophila
34	62.6	7.3	1152	12	BG309087	BG309087 HVSMEC000
35	62.4	7.2	928	17	CNS00DKY	AL071865 Drosophila
36	62.2	7.2	1125	9	AL547503	AL547503 Drosophila
37	62.2	7.2	1201	17	CNS016FX	AL106695 Drosophila
38	62	7.2	936	17	CNS07ANM	AL436808 T7 end of
39	62	7.2	958	17	CNS007AD	AL066801 Drosophila
40	62	7.2	1101	17	CNS00D77	AL075293 Drosophila
41	62	7.2	1101	17	CNS00YWL	AL096927 Drosophila
42	61.8	7.2	661	17	CNS020VJ	AL209800 Tetradon
43	61.8	7.2	836	17	AZ528131	AZ528131 ENTXY19TF
44	61.8	7.2	886	17	BH155984	BH155984 ENTPY20TF
45	61.8	7.2	957	17	BH154830	BH154830 ENTIT170TR

#### ALIGNMENTS

RESULT 1  
LOCUS CNS00L72 1101 bp DNA linear GSS 14-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL078714.1 GI:5102004  
VERSION AL078714  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

#### COMMENT

**THIS PAGE BLANK (USPTO)**



Fri Mar 28 13:43:34 2003

us-09-508-487-20.rng

Page 13

Job time : 246.586 secs

---

Db 4032 TTTTGTGTTTGGTAAATTTGGATGATGTTATATATATTTTGTAGTTATAGT 4091  
 QY 499 AAGCTCTGATAGAAATGCTCCAAAGCCGCTTTTAACTGGACCTGGGTAAGGAATGA 558  
 Db 4092 TTTTATTTATATTTTATTTTATGATATTTTATGATTTATTTATTTTATTTAT 4151  
 QY 559 TGTGGCAGGTGACAGTACTATGCGTGTGACAGATTTGACAGAAATTTATTTCCGTTA 618  
 Db 4152 AGTATTCAGATTTTAAATATTTTGTGTTAAATATATAGAGTTTGTTTTAAAGTAT 4211  
 QY 619 CATTTGCTAAATGATATATATAGAAACGAAATAGCCTTAATATAGCTTTGGAGGCT 678  
 Db 4212 AATTTATGCTAATATATATTTGGAATATTTAAATTTTAAACCGTTAAATATGAT 4271  
 QY 679 TTGAGCCAGTATTTGATATATATATAGGCGCAGGCTGCTGGCTTGAACATAGCTT 738  
 Db 4272 TTTTGTGTTTGTATTTTATTTAGTTGTGATGATAGAGATGTTGTATTTTAACTA 4331  
 QY 739 TCAGAAAAAGTATTTATTTTATTTTATTTTAAATGAGTAAAGCAATTTGTATTTG 798  
 Db 4332 AATGAGAGATTTTATATTAAGTAAATTTAGTTTATTTTGAATATTTAGTTTATTTT 4391  
 QY 799 GATTGCTGATTTGATTTGAAATTTAGAGCTTTGTTTATTTATTTATTTT 849  
 Db 4392 TATTTATTTATGATTTTGAATATTTATTTGATTTTGTGATTTT 4442

RESULT 15  
 AAS45410  
 ID AAS45410 standard; DNA; 11036 BP.

AC AAS45410;  
 DT 18-DEC-2001 (first entry)

XX Chemically pretreated genomic DNA associated with cell cycle #58.

XX Cell cycle: human; Cpg dinucleotide; cytosine methylation; HIV; aging;  
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;  
 KW graft-versus-host disease; glomerular disease; Levy body disease; cancer;  
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;  
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;  
 KW PCR primer.

XX Homo sapiens.

XX MO200168911-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02945.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602751/68.

XX Designing primers and probes for analysing diseases associated with

XX cytosine methylation state e.g. arthritis, cancer, aging,

XX PT arteriosclerosis comprising fragments of chemically modified genes

XX associated with cell cycle -

XX Claim 1; SEQ ID No 115; 28bp; English.

XX Sequences AAS45396-AAS45520 represent chemically pretreated genomic DNA

XX molecules associated with the cell cycle and specific PCR primers of the

CC invention. The sequences are useful for detecting the methylation state  
 CC of all Cpg dinucleotides in a sequence and therefore for analysing  
 CC associated diseases. By analysing cytosine methylations in the pretreated  
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy  
 CC of existing diseases or the predisposition to specific diseases can be  
 CC ascertained. The parameters may be compared to another set of genetic  
 CC and/or epigenetic parameters, the differences serving as basis for  
 CC diagnosis and/or prognosis events which are disadvantageous to patients.  
 CC The sequences of the invention are useful for the diagnosis and therapy  
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,  
 CC aging, glomerular disease, Lewy body disease, arthritis,  
 CC arteriosclerosis, solid tumours and cancers.

SO Sequence 11036 BP; 3049 A; 312 C; 2598 G; 5070 T; 7 other;

Query Match 7.6%; Score 65.2; DB 22; Length 11036;  
 Best Local Similarity 46.9%; Pred. No. 0.0049;  
 Matches 348; Conservative 0; Mismatches 378; Indels 16; Gaps 4;

QY 119 AGTAACTCTTGAAGAAAAAGTAAATTAATAGTTTAAAGAACCTTTTAAATTCAT 178  
 Db 6906 AGTGTATATTTTGTGTTTAAATTTATTTATTTATTTGCTATTTTATATATATTTAT 6965  
 QY 179 AATATGTTACTATTAATACCACTTTTAATTAAGAGCTTTTATGATTAATTTTATTTG 238  
 Db 6966 TGTATTTTGTAGGTTAAATTTATTTGATTTTGTGTTTGTGAGCTTTCGATTTAT 7025  
 QY 239 TGTATTTGCTAGCTTTTGTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 298  
 Db 7026 TAAACGATATTAATTTTACTGTTTGTGTTATGCTATATATATATATATATATATAT 7085  
 QY 299 TTTTAATTTGGAGCGGAGAAAAAGTAAATTTAGCTTATGCAAGTAAAGAAAGATCCAT 358  
 Db 7086 TTTGATATATTTTGTGTTTGAATAGTAT---TTATAGATTTAGATTTAGTATATAT 7141  
 QY 359 TGTGCAATTTTATTTGAACCTTTTATAGGCTTTGGAATAGCTTCTTGTCTCAAGAGA 418  
 Db 7142 TTTATATATGTTATGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 7201  
 QY 419 TATCT-----TGCGGCTTTTCTTATTTGCTGATGAGCTGCTGATAGGCT 468  
 Db 7202 GATTTTATGTAAT 7261  
 QY 469 TAAATCTTACAGAGCTTATTTAGATTAATCAAGCTTGTATAGAAATGCTCAAGACCG 528  
 Db 7262 ATTTAGTTTATGATTTATTTATGATTAATTTATTTATTTATTTATTTATTTATTTAT 7321  
 QY 529 CTTTATGCTGACCTTGGGCTAAGGA-ATGATGTTGGCAGCTGCACTTACTAGCTGTG 587  
 Db 7322 TTTGTTTTTTTATTTTAAATGAGAGTGTGTTGTAATTAATGAATATGATGATG 7381  
 QY 588 ACAAGATTTGACAGAAATTTATTTCCGTTTACATTTGCTATTTAGTTAAATAGAACTG 647  
 Db 7382 AATAGTTATTTTAAATGTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7441  
 QY 648 AAAATAGCCTTAATATATAGCTTTTGAAGGCTTGAAGCTTGTGATTTATATATGCGC 707  
 Db 7442 AGAAGTTTGTGTTGAGGCTGCGAAATATATAGTTAGTAAATGTTTGTGATATTTT 7501  
 QY 708 CAAGCTAGCGCTTTGGCTTGAACATCTTTCAAAAAAGTTATTTATTTATTTATTTAT 767  
 Db 7502 TTTTATTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7560  
 QY 768 ATTTAAATGAGTATAGCAATTTTGTATTTGATTTGCTATTTGATTTGAAGAAATAGAGC 827  
 Db 7561 ATTTTAAATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7620  
 QY 828 TTTTGTATTTATTTATTTATTTT 849  
 Db 7621 TTTTGTATTTATTTATTTATTTT 7642

Search completed: March 27, 2003, 01:32:16

OY	51	TTATTAAAGCAATGGAATTCCTAATTTTAATGTGCACAAATCAAGAAGCCTATGGGAAGC	110		
Db	3973	TAAATATATATATATTTTAAATAGGAGAAATTCGTTAAAATAAGGGGTATATGATTTTATG	40322		
OY	111	GAAATTCAGTAATATACTTTGAAAAAGTAAATTTAAATAGTTTTTAAAAACCCTTTTNA	170		
Db	4033	TAAATTTGAATTTAATGAGGTAGTATTTAAATTTTAAATTTTAAATATTTTGTG	40922		
OY	171	ATTTCATTAATATGTACTATATACCAGTTTAAATAAAGAGGTTTTATGA-----A	223		
Db	4093	AATTTATGTTTTTATATTTTAGGTTATATATGATGAATGAAGTGGGTTTTATGGTTTTTGA	41522		
OY	224	TAAATTTTATATGTCTTTTGTCCATCCCTTTGTCTTTGTCTTTGTCTTTGTCCATCGCA	283		
Db	4153	TATTTTGTCTTTTGTCTTTTGTCTTTGTCTTTGTCTTTGTCTTTGTCTTTGTCTTTGTCTTT	42122		
OY	284	TGATTCATAAAGCCCTTTTAAATTTGGAGCGGAGAAAAACCTTTAGCTTATGAACATAG	343		
Db	4213	TATATGAGTGTGTAGTTTTTTTAAAGTATAGTATAGTATCTCGGTAGAAATTAATTT	42722		
OY	344	TACAAAGATCTTATTTGCCATTTTNTAATGAACTTTTNTAAGGTTTGGAAATAGTTTC	403		
Db	4273	TAGGAGCAGAGATGCTGTTTTTTTTTTTGTAGTTTTATAGTAGTGTATAGGAGGA	43322		
OY	404	TTTTGCTCAGAGAAATATCTTGGGGTTTTCTATCTGTGATTTGATGCACTTGCTAT	463		
Db	4333	TTTTGTGTGGGGTTTTTAAATTTTATTTTATATATATATATATATATATATATATATAT	43922		
OY	464	AGGTTTATATCTTACAGACCTTATTTAGATATCAAAAGCTCTGATGAAGATGCTCCAA	523		
Db	4393	TTAATTTTTTTTTTTTTTTTTTTTATATAGATTTTAAATATATTTTTTTTTTAACTTTAAAT	44522		
OY	524	AGCCGCTTTTAACTGACTTGGGTTAAGGAATGATGTTGGCAGGTGCAGTTACTATGTC	583		
Db	4453	TTTATAGATTTTAAAGGAAGGCTAAATATGTTATAGTTTTTTTGTATAGTATATAG	45122		
OY	584	TGTACACAAGTGTACAGAAATTTATTTCCGTTTACATTTGCTAATAGTATATATAGCAA	643		
Db	4513	--GATTTATTTTATTTTAACTTTTATATAGTTTTTAAATTTTATTTTAAAGTATATATAG	4569		
OY	644	ACTGAAATAATAGCCTTAAATATACCTTTTGAGAGGTTTGAGCCTAGCTTTGATATTAAT	703		
Db	4570	TGTGATTTTATTTTATTTTATTTATTTATTTATTTATTTTGGTTAAATATTTTAAAGAGTTT	4629		
OY	704	GGGCCAAGCTAGCGCTCTTGGGTTGAACTATCTTCAAAAAAGTATATTAATTTATTT	763		
Db	4630	TAGAAATTTTAAATTTTGTGATTTTATTTTATTTTGTAGTTTGTATATTTTAAATTT	4689		
OY	764	TATATATTAATAGTATAGCAATTTTGTATTTGATTTGCTCATTTGATTTGAAAATTA	823		
Db	4690	TTGTTTCTTTTATTTAGTTTAAAGCTTTTNTTAAATTTTGGTATTTTATAGTAAATTT	4749		
OY	824	GAGCTTTGTTTATATTTATTTATTTTATTTTCTCTGCTAA	862		
Db	4750	TTATTTTGTATTAATTTTATTTGATTTAGTTTATTTTAA	4788		
<hr/>					
RESULT_14					
ID	ABL70480	standard; DNA; 9742 bp.			
XX	ABL70480;				
DT	01-JUL-2002	(first entry)			
DE	Chemically treated cell signalling DNA sequence complementary to#185.				
KM	Cell signalling: cytosine methylation; cell signalling disease;				
XN	cancer; tumour; cytostatic; ds.				
OS	Unidentified.				
FN	W0200202807-A2.				

XX	PD	10-JAN-2002.
XX	PF	29-JUN-2001; 2001WO-EP07471.
XX	PR	30-JUN-2000; 2000DE-1032529.
XX	PR	01-SEP-2000; 2000DE-1043826.
XX	PA	(EPIG-) EPIGENOMICS AG.
XX	PI	Olek A, Piepenbrock C, Berlin K;
XX	DR	WPI; 2002-154758/20.
XX	PT	Nucleic acid, useful for diagnosis and therapy of diseases associated with cell signalling e.g. cancer, comprises chemically modified genomic sequences of genes associated with cell signalling -
PS	PS	Claim 1; SEQ ID NO 370; 24ppt-sequence listing; English.
CC	CC	The invention relates to a nucleic acid comprising a sequence of at least 18 bases of a segment of chemically pretreated DNA of genes associated with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or PNA-oligomers for detecting cytosine methylations, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records AB170111-AB170626 represent chemically pre-treated genomic DNA's of genes associated with cell signalling. CC Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.
SQ	SQ	Sequence 9742 BP; 2945 A; 88 C; 1819 G; 4890 T; 0 other;  Query Match            7.6%; Score 65.8; DB 24; Length 9742; Best Local Similarity   44.0%; Pred. No. 0.0038; Matches   366; Conservative   0; Mismatches   462; Indels   3; Gaps   2
QY	QY	19 ACAATTTGTAGCAATTTTCGCACAAATAAAGACATTATAAAGAATTGAATAATTCCTAATTT 78
DB	DB	3615 ATAAGTGAAAAAATTTGTAATGATATATAATTAATTTGAATACGTAATTAATAATTTTAA 3674
QY	QY	79 TATGGTCAAATCACAAGACTCTATTGGGCAAGCGAATTTCAAGTAATCTTGAAAAAAGT 138
DB	DB	3675 TTATAATAAAATTTTGAAAAT-AAATTTTAGCAATTTTATATTTTATTAAAGTATATATACA 3733
QY	QY	139 TAAATTTAAATAGTTTTTAAAAACCCTTTTAAATTTCAATATATGTTCTPATAATACGA 198
DB	DB	3734 AAATATGTAATTBGGTTAATAGAAAATATTTTTTAAAAAATTGAAAAAGTAA--ATAATAGTA 3791
QY	QY	199 GTTTTAATAAAGAGCTTTTATGAAATAATTTTAAATGTTGTTTGGCTAGCCCTTTTGCG 258
DB	DB	3792 ATTTTTTTTATATGTTTTTTTTTTTATTTTATTTTCTTAAAGAAAAAATAA 3851
QY	QY	259 TTTTTCCTAGCTTTGCTCAAGCTGATGATCTTAAAGCCCTTTAATTTGGAGCCGGAG 318
DB	DB	3852 TTATTAATTTTATAGAGGGTAGAGAGCTAATTTTAAATGATTTGAAATTTTGTGCTAGTA 3911
QY	QY	319 AAAAACCTTTAGCTATAGAACTAGTAAAGAAAGATCCATTTGCCATTTTATGCAAC 378
DB	DB	3912 TTCGGTGTAGATGAGGTTAAAGTAAAGGATTAAATATTTATTTGAAAAATATTTAGATATT 3971
QY	QY	379 TTTTATTTTATGAGTTTGAATAGGTTCTTTTGTCTCAAGAGATATTCCTWGGSCTTTCTTA 438
DB	DB	3972 ATGCTTTGAGTAATATGCTTTTGTAATTTGAAAAAATTTTTTATTTTGTGTATATTTG 4031
QY	QY	439 TTTCTGGATTTTGATCGAGTTGGTATAGGGTTAATCTTACAGGACCTTATTTAGATATCA 498





```

      Cons: 337; Mismatches 0; Indels 8; Gaps 1;
      Del: 0; Substitutionally 49.1%; Freq: NO: 0.0022;

```

PD 20-SEP-2001.  
YY

	2	C	Z	.
	O	E	S	
	X			

KW	Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KM	Epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW	Infection; diagnosis; characterisation; detection; ds.
XX	
OS	Borrelia burgdorferi.
PN	
XX	WN9856943-A1.
XX	
PD	30-DEC-1998.
XX	
PF	18-JUN-1998; 98WO-US12764.
XX	
PR	03-SEP-1997; 97US-0057483.
PR	20-JUN-1997; 97US-0050359.
PR	22-JUL-1997; 97US-0053344.
PR	22-JUL-1997; 97US-0053377.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(MEDT-) MEDIMMUNE INC.
PI	Clayton R, Dougherty BA, Fraser C, Iathigra R, Smith HO;
DR	White OR;
XX	
PT	New isolated Borrelia burgdorferi nucleic acids - used to develop
PT	products for the detection, diagnosis, characterisation, prevention
PT	and therapy of infections, particularly Lyme disease
XX	
PS	Claim 1; Page 925-931; 1128bp; English.
XX	
CC	AAX20248 to AAX20402 represent polynucleotide sequences isolated from
CC	Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC	the detection, diagnosis, characterisation, prevention and therapy of
CC	Bb infections, e.g. Lyme disease. They can also be used for the
CC	production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC	to a family of motile, spiral-shaped bacteria called Spirochetes.
CC	Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC	endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX	Lyme disease.
XX	
SO	Sequence 9399 BP; 3470 A; 1281 C; 1049 G; 3595 T; 4 other:
	Query Match            10.2%; Score 88.2; DB 20; Length 9399;
	Best Local Similarity   51.9%; Pred. No. 3.5e-07;
	Matches 229; Conservative 0; Mismatches 203; Indels     9; Gaps
OY	305 TTTGGGACCGGAGAAAACTTTTAGCTTATGAACCTAGTAAGAAAAGATCCTATTGTGCC 364
b	
OY	8061 TGTGTGAAATATTCGCACCCTAATGAAAATGMAAAAGCAACCATATTAGCACCC 8120
b	
OY	365 ATTTTATTTGAACCTTTTTTTAGGCTTGGAATAGGTTCTTTGGCAAGAGATATTCT 424
b	
Db	8121 ATTACTTTTGAACAATCTTTTATCTTTAGAAATAGAGATCTTTTGTCCAAGGAGATTATAT 8180
OY	425 TGGGGTTTCTTATCTTGGAATTTGATGACAGTTGGTATTAAGGTTAATCTTACAGAAC 484
b	
Db	8181 TGTGTGTGTGCACATCTTGGCTCTCAAGTGTGGAGAAATACATAATTAATGGCTGG--- 8237
OY	485 TTATTTAATATCAAAAGCTCTTGATTAAGAATCTCCAAAAGCCGCTTTTAACTGACCTGG 544
b	
Db	8238 -----ATACATAGACCGGGGGAATTTGGATTCTTTCACAGAAACACAGCAACGTATAAC 8291
OY	545 GGGTAAAGGAATGATGTTGGCAGGTGCATGTTCTTGGCTGTGACAAAGTTACAGAAAT 604
b	
Db	8292 CGGAGAGATATTATACGAATATGAGAGGGGCTCCAAATTTGCAGCATCTTACATACGTGAGAT 8351
OY	605 TATTATTCCTGTTACATTGCTTAATAGTTATATATAGAAACTGAATAATAGCCTTAATAT 664
b	
Db	8352 TATCATTCATTAATAATTTTGTAATAGATPACAAAGCAGACCTTTAAAAAAAAGACTGGCAT 8411
OY	665 AGCTTTTGGAGGGTTGAGCCTAGTTTGTGATTTATATATGGGCCAAAGCTAGCGCTCTGGG 724
b	

[illegible]





XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (MEDI-) MEDIMUNE INC.  
 XX ChOI GH, Erwin AL, Hanson MS, Lathigra R;  
 XX WPI: 1999-189980/16.  
 DR P-PSDB: AAY19809.  
 XX  
 XX New isolated *Borrelia burgdorferi* nucleic acids - used to develop  
 PT products for the diagnosis, prevention and treatment of diseases  
 PT caused by *Borrelia*, particularly Lyme disease  
 XX  
 XX Claim 1; Page 73; 275pp; English.  
 CC This sequence encodes a *Borrelia burgdorferi* (Bb) protein of the  
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides  
 CC can be used in vaccines for eliciting protective antibodies to members of  
 CC the *Borrelia* genus, particularly for the use against Lyme disease in  
 CC humans and animals. They can be used for preventing or attenuating an  
 CC infection caused by a member of the *Borrelia* genus. The products can also  
 CC be used for detection of members of the *Borrelia* genus.  
 SQ  
 Query Match 42.9%; Score 369.4; DB 20; Length 477;  
 Best Local Similarity 86.1%; Pred. No. 7.5e-58;  
 Matches 409; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
 OY 281 TGAATGTTTAAAGCGCTTTTAAATTTGGAGCGGAGAAAACTTTAGCTTATGAAC 340  
 DB 3 TAAATCTAAAGATGCTGGTGGATGAGTGTGAGAAAACTTTGGTATGAAAC 62  
 OY 341 TGTAGAAGATCCATTTGTCATTTTATGAACCTTTTATGGCTTGGATAGG 400  
 DB 63 TAGCAAGCAAGATCTATTGTACCATTTTATGAACCTTTTATGGCTTGGATAGG 122  
 OY 401 TTTCTTTGCTCAAGAGATATCTTGGGGGTTTCTTATTTCTGATTTGATGACAGTGG 460  
 DB 123 CTCCTTTGCTCAAGAGATATCTTGGAGTCTCTATTTCTGATTTGATGACAGTGG 182  
 OY 461 TTTAGGTTAATCTTACAGAGCTTATTTAGATCAAAAGCTTTGATPAAGATGCTCC 520  
 DB 183 TTTAGGCTTAAATCTTACAGAGCTTATTTAGATCAAAAGCTTTGATPAAGATGCTCC 520  
 OY 521 AAAAGCGCTTTTAAATGAGCTTTGGGTAAGGAATGATGGGAGCTTGAATCTTAA 580  
 DB 243 AAAAGCTGTTTCAATGAGCTTTGGGTAAGGAATGATGGGAGCTTGAATCTTAA 580  
 OY 581 GGCTGTGACAGATGACAGAAATTTATTCCTTACATTTGCTAATAGTTAATAG 640  
 DB 303 GGCTGTGACAGATGACAGAAATTTATTCCTTACATTTGCTAATAGTTAATAG 640  
 OY 641 GAAATGAAAAATAGCTTAAATATAGCTTTGGAGGCTTGGACCTAGTTGATTTAA 700  
 DB 363 GAGCTAAAAAATAGCTTAAATATAGCTTTGGAGGCTTGGACCTAGTTGATTTAA 700  
 OY 701 TATGGCCCAAGCTAGGCTTGGGTTTGAACATCTTCAAAAAAGTTATTA 755  
 DB 423 AATGGCCCAATGAGCTTGGGTTTGAACATCTTCAAAAAAGTTATTA 755  
 RESULT 6  
 AAX20266  
 ID AAX20266 standard; DNA; 7074 BP.  
 XX  
 XX AAX20266;  
 AC  
 XX  
 XX 04-MAY-1999 (first entry)  
 DT  
 XX  
 DE *Borrelia burgdorferi* polynucleotide sequence #19.  
 XX  
 XX *Borrelia burgdorferi*; spirochete; bacterium; pathogen; Lyme disease;

KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
 KW infection; diagnosis; characterisation; detection; ds.  
 XX  
 XX *Borrelia burgdorferi*.  
 OS  
 XX WO9858943-A1.  
 PN  
 XX 30-DEC-1998.  
 PD  
 XX  
 XX 18-JUN-1998; 98WO-US12764.  
 PF  
 XX  
 XX 03-SEP-1997; 97US-0057483.  
 PR 20-JUN-1997; 97US-0050359.  
 PR 22-JUL-1997; 97US-0053344.  
 PR 22-JUL-1997; 97US-0053377.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (MEDI-) MEDIMUNE INC.  
 XX Claydon R, Dougherty BA, Fraser C, Lathigra R, Smith HO;  
 PI White OR;  
 PI WPI: 1999-081217/07.  
 DR  
 XX  
 XX New isolated *Borrelia burgdorferi* nucleic acids - used to develop  
 PT products for the detection, diagnosis, characterisation, prevention  
 PT and therapy of infections, particularly Lyme disease  
 XX  
 XX Claim 1; Page 951-955; 1128pp; English.  
 XX  
 XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from  
 CC *Borrelia burgdorferi* (Bb). Products derived from Bb can be used for  
 CC the detection, diagnosis, characterisation, prevention and therapy of  
 CC Bb infections, e.g. Lyme disease. They can also be used for the  
 CC production of biosynthetic products, e.g. enzymes. *Borrelia* belongs  
 CC to a family of mobile, spiral-shaped bacteria called Spirochetes.  
 CC Spirochetes are pathogenic in humans and *Borrelia* causes epidemic and  
 CC endemic relapsing fever, and Lyme borreliosis, more commonly known as  
 CC Lyme disease.  
 SQ  
 Query Match 11.6%; Score 100; DB 20; Length 7074;  
 Best Local Similarity 50.5%; Pred. No. 2.6e-09;  
 Matches 410; Conservative 0; Mismatches 355; Indels 47; Gaps 5;  
 OY 39 CAATTAAGACATTATTAAGAAATGAAATGCTAATTTATGTCATAAGAGCT 98  
 DB 2982 CTAAATAGAGAGCTCTATTGATTCATCAAAAAAGTATTTAAATCAAAATCACTAT 3041  
 OY 99 CTATTGGGAAGCAATTTCAAGTAATCTTGAAGAAAGTTAATTAATTAAGT-TTAA 157  
 DB 3042 TTTATTAACAAATCATCTTAACCCAAATTTCAATAGAAATTAATGATTTCTTGG 3101  
 OY 158 AAACCTTTTAAATCTTAAATATGTTACTTAATTAATACAGTTTAAATGAAGAGCTTT 217  
 DB 3102 GGGTTTTATTAATGTTGTTGCTAGCCTTTGCTGTTTCTTACCTTGGCTCA 3161  
 OY 218 TATGAATTAATTTAATGTTGTTGCTAGCCTTTGCTGTTTCTTACCTTGGCTCA 3161  
 DB 3162 TATGAATTAATTTAATGTTGTTGCTAGCCTTTGCTGTTTCTTACCTTGGCTCA 3221  
 OY 278 AGCTGATGATTTCAAAAGGCTTTTAAATTTGGAGCGGAGAAAACTTTAGCTTATCA 337  
 DB 3222 ACAAGATTAAGCTTGAAGAAAGTGT- - - - -GGAAGTATTTGAACCATTAAGAAATTA 3275  
 OY 338 AACTAGTAAGAAAGATCTATTGTCATTTTATGGAACCTTTTATGGCTTGGAAAT 397  
 DB 3276 AAGCGAAAAAGCACTATCTAGACCATCTTTTGAATTTATTTTAACTTTAGGAAT 3335  
 OY 398 AGGTTCTTTGCTCAAGAGATATCTTGGGGGTTTCTTATTTCTTGAATTTGATGACAGT 457  
 DB 3336 AGGATCTTTGTCGAAGAGATTAATTTGTTGGTGGCCAGAGCTTGATCTCAGTTAAT 3395

```

Db 301 GCAAGATCCTATGTACCAATTTATATGACCTTTTATAGGTTGGAATAGGCTCCT 360
Qy 407 TGTCTAAGAGATATTTCTGGGGTTTCTTATTTCTTGATTTGATGACGTTGATAGG 466
Db 361 TGTCTAAGAGATATTTCTGGGGTTTCTTATTTCTTGATTTGATGACGTTGATAGG 420
Qy 467 GTTATATCTTACAGAGCTTATTTATGATATCAAGCTCTTGATTAAGATGCTCCAAAGC 526
Db 421 GCTTATATCTTACAGAGCTTATTTATGATATCAAGCTCTTGATTTGATGACGTTGATAGG 480
Qy 527 CGCTTTAAGTACTTGGGGTGAAGGATGATGTTGGCAGTGCATCTATGCTGT 586
Db 481 TGTCTTCAAGTACTTGGGGTGAAGGATGATGTTGGCAGTGCATCTATGCTGT 540
Qy 587 GACAAGATGACAGAAATTTATTTCCGTTTACATTTCTTATAGTTATATAGAACT 646
Db 541 GACAAGATGACAGAAATTTATTTCCGTTTACATTTCTTATAGTTATATAGAACT 600
Qy 647 GAAAAATAGCCTTATATATAGCTTTTGAAGGTTGAGCCTTATGATATATATAGG 706
Db 601 AAAAAATAGCCTTATATATAGCTTTTGAAGGTTGAGCCTTATGATATATAGG 660
Qy 707 CCAAGCTAGCCTCTTGGGTTTGAACATATCTTTCAAAAAAGTTATATTTATTTAT 766
Db 661 CCAATCCAGTCTCTTGGGTTTGAACATCTTTCAAAAAAGTTATATTTATTTAT 719
Qy 767 TATTAATGAGTGTATGACATTTGTATGTGATGCT 805
Db 720 TACAAAAATGGGTATGTCATTCGTATGAAATGGT 758

```

RESULT 4  
AAx61505  
ID AAX61505 standard; DNA; 540 BP.

AAx61505:

19-JUL-1999 (first entry)

B. burgdorferi antigenic protein coding sequence, f933.nt.

Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

Borrelia burgdorferi.

MO9859071-A1.

30-DEC-1998.

18-JUN-1998; 98WO-US12718.

03-SEP-1997; 97US-0057483.

20-JUN-1997; 97US-0050359.

22-JUL-1997; 97US-0053344.

22-JUL-1997; 97US-0053377.

(HUMA-) HUMAN GENOME SCI INC.

(MEDI-) MEDIMUNE INC.

Choi GH, Erwin AL, Hanson MS, Lathigra R;

WPI; 1999-189980/16.

P-PSDB; AAY19808.

New isolated Borrelia burgdorferi nucleic acids - used to develop

products for the diagnosis, prevention and treatment of diseases

caused by Borrelia, particularly Lyme disease

Claim 1; Page 73; 275pp; English.

This sequence encodes a Borrelia burgdorferi (Bb) protein of the

invention, which is suitable for use in a vaccine. The Bb polypeptides

CC can be used in vaccines for eliciting protective antibodies to members of  
CC the Borrelia genus, particularly for the use against Lyme disease in  
CC humans and animals. They can be used for preventing or attenuating an  
CC infection caused by a member of the Borrelia genus. The products can also  
CC be used for detection of members of the Borrelia genus.

SO Sequence 540 BP; 145 A; 71 C; 123 G; 201 T; 0 other;

Query Match 48.7%; Score 420; DB 20; Length 540;

Best Local Similarity 87.4%; Pred. No. 5.9e-67;

Matches 472; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

```

Qy 219 ATGATAAATTTTAAATGTTGTTTGGTACCTTTGTTGTTTCTTATGCTTGCCTCA 278
Db 1 ATGATAAATTTTAAATGTTGTTTGGCAACCTTTGTTTCTTATGCTTGCCTCA 60
Qy 279 GCTGATGTTCTAAA--GCGCTTTATTTGGGAGCGGAGAAAACCTTTAGCTTAT 335
Db 61 GCTATGATTTCTAAAATGATGCGCTTGGGATGAGTCTGAGAAAACCTTTGCTTAT 120
Qy 336 GAACTAGTAAAGAGATCTATTTGCTATTTTATGAACTTTTATGAGGTTTGA 395
Db 121 GAACTAGTAAAGAGATCTATTTGCTATTTTATGAACTTTTATGAGGTTTGA 180
Qy 396 ATAGGTTCTTTGCTCAAGAGATATCTTGGGGTTTCTTATTTCTGATTTGATGCA 455
Db 181 ATAGGTTCTTTGCTCAAGAGATATCTTGGGGTTTCTTATTTCTGATTTGATGCG 240
Qy 456 GTTGTATAGGTTTATATCTTACAGAGCTTATTTAGATATCAAGCTCTTGTATAGAT 515
Db 241 GTTGTATAGGTTTATATCTTACAGAGCTTATTTAGATATCAAGCTCTTGTATAG 300
Qy 516 GCTCCAAAAGCGCTTTAAGTGGACTTGGGGTGAAGGATGATGAGGTCAGTGCAGT 575
Db 301 ACTAAAAGAGCTCTTTTCAATGACCTTGGGTTAAGGAGTTATGTCAGGTTGCTT 360
Qy 576 ACTATGCTGTGACAGATTTGACAGAAATTTATTTCCGTTTACATTTGCTAATAGTTAT 635
Db 361 ACTATGCTGTGACAGATTTGACAGAAATTTATTTCCGTTTACATTTGCTAATAGTTAT 420
Qy 636 AATAGGAACCTGAAAATAGCTTATATTTAGCTTTTGGAGGTTTGAGCCTTATTTGAT 695
Db 421 AATAGGAACCTGAAAATAGCTTATATTTAGCTTTTGGAGGATTTGAACTTATTTGAT 480
Qy 696 ATTATATGAGGCAAGCTAGGCTCTTGGGTTTGAACATCTTTCAAAAAAGTTATTA 755
Db 481 GTTGCAATGGGCCAATCCAGTGTCTTGGTTTGAACATCTTTCAAAAAAGCTATTTAA 540

```

RESULT 5  
AAx61506  
ID AAX61506 standard; DNA; 477 BP.

AAx61506:

19-JUL-1999 (first entry)

B. burgdorferi antigenic protein coding sequence, t933.nt.

Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

Borrelia burgdorferi.

MO9859071-A1.

30-DEC-1998.

18-JUN-1998; 98WO-US12718.

03-SEP-1997; 97US-0057483.

20-JUN-1997; 97US-0050359.

22-JUL-1997; 97US-0053344.

22-JUL-1997; 97US-0053377.

CC a more specific and sensitive antibody response, and diagnosis,  
 CC compared with use of whole bacteria as antigen.  
 XX  
 XX  
 SQ Sequence 749 BP; 234 A; 82 C; 150 G; 283 T; 0 other;

Query Match 70.4%; Score 607; DB 20; Length 749;  
 Best Local Similarity 90.1%; Pred. No. 1.3e-100;  
 Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

OY 27 TGAATTTTCGCAATTAAGACATATTAAGAAATGAAATGCTAATTTATGCTCA 86  
 DB 1 TGAATTTTCACAAATTAAGATATTTGTTAAAGAAATGAAATGCTAATTTATGCTCA 60  
 OY 87 AATCAAGAAGCTCTATTTGGGAAGCGAATTCAGATATCTATTTGAAAAAGTAAATTTA 146  
 DB 61 AATCAAGAAGCTCTATTTGGGAAGCGAATTCAGATATCTATTTGAAAAAGTAAATTTA 120  
 OY 147 AATAGTTTAAACCTTTTAAATTTCAATTAATATGTTACTATAATACCACTTTTAAT 206  
 DB 121 AATAGTTTCAAAAGCCCTTTTAAATTTCAATTAATATGTTACTATAATACCACTTTTAAT 179  
 OY 207 AAGAGGTTTTTATGAAATTTTAAATTTGTTTGTACCTTTTGTCTTTTCTTCT 266  
 DB 180 AAGAGGTTTTTATGAAATTTTAAATTTTAAATTTTGTATATCTTTTGTCTTTTCT 239  
 OY 267 AGCTTTGCTCAAGCTGATTTCTAAAGCGCTTTTAAATTTGGAGCGGAGAAAAACTT 326  
 DB 240 AGTTTGTCTCA--GATATTTCTAAAGACTTTTAACTGTGGACGGAGAAAAATTT 296  
 OY 327 TTAGCTTAAGAACTAGTAAGAAAGATCTTATTTGCTCATTTTAAATTTAAACCTTTTAA 386  
 DB 297 TTGCTTAAAGAACTAGTAAGAAAGATCTTATCTTCAATTTTAAATTTAAACCTTTTAA 356  
 OY 387 GGGTTTGAATAGGTTCTTTTGTCTCAAGGATATCTTGGGGCTTTTCTTATCTTGA 446  
 DB 357 GGGTTTGGATAGGTTCTTTTGTCTCAAGGATATCTTGGAGGCTTCTTATCTTGA 416  
 OY 447 TTTGATGCAATGCTTATGAGGTTAATCTTACAGAGCTTATTTAATATCAAGCTCTT 506  
 DB 417 TTTGATGCGGTTGCTTATGAGGTTAATCTTACAGAGCTTATTTGACATCAAGATTTT 476  
 OY 507 GATAGAAATGCTCAAAAGCCCTTTTAAAGTGACCTTGGGTAAGGAATGATGTTGCA 566  
 DB 477 GATAGAAATGCTCAAAAGCCCTTATTTAAAGTGACCTTGGGTAAGGAATGATGTTGCA 536  
 OY 567 GGTGCAATGCTTATGAGGCTTGAACATGACGAATATTTTCTTCCGTTTACATTTGCT 626  
 DB 537 GGTGCTTACTATGCTGCTGACAAATGCTGACGAATATGCTTCCATTTTACATTTGCT 596  
 OY 627 AATAGTTTAAATAGGAAGCTGAAATATAGCTTATATAGCTTTTGAGGCTTTGAGCT 686  
 DB 597 AATAGTTTAAATAGGAAGCTGAAATATAGCTTATATAGCTTTTGAGGCTTTGAGCT 656  
 OY 687 AGTTTGAATTAATATGAGGCAAGCTAGCGCTTTGGTTTGAACATCTTTCAAAAA 746  
 DB 657 AGTTTGAATTAATATGAGGCAAGCTAGCGCTTTGGTTTGAACATCTTTCAAAAA 716  
 OY 747 AGTTTGAATTAATATGAGGCAAGCTAGCGCTTTGGTTTGAACATCTTTCAAAAA 773  
 DB 717 AGCTATTAATTTTATTTATCTAGAAAA 743

RESULT 3  
 AAX30098  
 ID AAX30098 standard; DNA; 759 BP.  
 AAX30098;  
 XX  
 AC 17-JUN-1999 (first entry)  
 XX  
 DT  
 XX  
 DE Borrelia burgdorferi B31 protein encoding cDNA.  
 XX  
 KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.

XX OS Borrelia burgdorferi.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 170..709  
 XX FT /tag=a  
 XX PN W0912960-A2.  
 XX PD 18-MAR-1999.  
 XX PF 04-SEP-1998; 98WO-1B01424.  
 XX PR 16-SEP-1997; 97US-0059036.  
 XX PR 10-SEP-1997; 97DK-0001041.  
 XX PA (SYMB-) SYMBICOM AB.  
 XX PI Bergstrom S;  
 XX DR WPI: 1999-215027/18.  
 XX DR E-PSDB; AA104278.  
 XX PT Nucleic acid from Borrelia burgdorferi encoding virulence associated protein P13  
 XX PS Claim 1: Page 107-108; 118pp; English.  
 CC The present sequence encodes a Borrelia burgdorferi B31 protein. The  
 CC present invention describes an isolated nucleic acid (I) that: (i)  
 CC encodes a polypeptide fragment (II) immunologically reactive with  
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of  
 CC Borrelia burgdorferi B31, but not reactive with most proteins from  
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions  
 CC to specific nucleic acid sequences but not to genomic DNA from most  
 CC other spirochaetes. (I), (II) and transformed cells are useful in  
 CC vaccines to protect against infection by B. burgdorferi sensu lato.  
 CC (I), (II) and antibodies raised against (II) are used to diagnose such  
 CC infections. In standard immunoassays or amplification/hybridization  
 CC tests, (I) are also used to produce recombinant (II). The 13 kD protein  
 CC is involved in virulence and is highly conserved within B. burgdorferi  
 CC sensu lato, but is absent from other Borrelia species (e.g. those  
 CC responsible for relapsing fever or avian borreliosis). It should provide  
 CC a more specific and sensitive antibody response, and diagnosis,  
 CC compared with use of whole bacteria as antigen.  
 XX  
 SQ Sequence 759 BP; 228 A; 90 C; 157 G; 284 T; 0 other;  
 Query Match 69.2%; Score 596.6; DB 20; Length 759;  
 Best Local Similarity 88.4%; Pred. No. 1e-98;  
 Matches 671; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

OY 50 ATTATTAAAGAAATGAAATGCTAATTTTATGTCAAATCAAGAGCTCTATGGAAG 109  
 DB 1 ATTGTTAAAGAAATGAAATGCTAATTTTATGTCAAATCAAGAGCTCTATGGAAG 60  
 OY 110 CGAATTTCAAGTAATCTTTGAAAAAGTAAATTTAAATAGTTTAAAAACCTTTTAA 169  
 DB 61 CGAATTTCAAGTAATCTTTGAAAAAGTAAATTTAAATAGTTTAAAAACCTTTTAA 120  
 OY 170 AATTTCAATTAATGTTACTATATACAGCTTTTAAATTAAGAGCTTTTATGAATTAAT 229  
 DB 121 AATTTCAATTAATGTTACTATATACAGCTTTTAAATTAAGAGCTTTTATGAATTAAT 180  
 OY 230 TTTAATTTGTTTGTCTAGCCCTTTTGTGTTTCTAGCTTTGCTCAAGCTGATGATTC 289  
 DB 181 TTTAATTTTGTGTTGGAACCTTTTGTGTTTCTAGCTTTGCTCAAGCTGATGATTC 240  
 OY 290 TAAAG--GCGCTTTTAATTTGGAGCGGAGAAAAACTTTTATGCTTAAGAACTAGTAA 346  
 DB 241 TAAAGATGCGCTTTGGAGTGAAGTGAAGAAAACTTTTATGAACTAGTAACTAGCAA 300  
 OY 347 GAAAGATGCTATTTGCTCAATTTTATTTGAACCTTTTATGAGGCTTGAATAGTCTTT 406

DR P-PSDB: AAY04279.  
 XX Nucleic acid from *Borrelia burgdorferi* encoding virulence associated  
 PT protein p13  
 PS Claim 1; Page 110-111; 118pp; English.  
 XX  
 CC The present sequence encodes a *Borrelia afzelii* ACAI protein. The  
 CC present invention describes an isolated nucleic acid (I) that: (1)  
 CC encodes a polypeptide fragment (II) immunologically reactive with  
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of  
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from  
 CC other spirochaetes; and/or (II) hybridizes under stringent conditions  
 CC to specific nucleic acid sequences but not to genomic DNA from most  
 CC other spirochaetes. (I), (II) and transformed cells are useful in  
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.  
 CC (I), (II) and antibodies raised against (II) are used to diagnose such  
 CC infections, in standard immunoassays or amplification/hybridization  
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein  
 CC is involved in virulence and is highly conserved within *B. burgdorferi*  
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those  
 CC responsible for relapsing fever or avian borreliosis). It should provide  
 CC a more specific and sensitive antibody response, and diagnosis,  
 CC compared with use of whole bacteria as antigen.  
 CC  
 XX  
 SQ Sequence 862 BP; 265 A; 94 C; 162 G; 341 T; 0 other;  
 Query Match 100.0%; Score 862; DB 20; Length 862;  
 Best Local Similarity 100.0%; Pident. No. 1.8e-146;  
 Matches 862; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GATTTCATTCGATCCGAGATTTGTAGAAATTTTCGACAAATAAGACATTATTAAG 60  
 DB 1 GATTTCATTCGATCCGAGATTTGTAGAAATTTTCGACAAATAAGACATTATTAAG 60  
 QY 61 AATTGAATTCGCTAATTTATGCTCAAAATCAAGAACTCTATTGGGAACGCAATTCAAG 120  
 DB 61 AATTGAATTCGCTAATTTATGCTCAAAATCAAGAACTCTATTGGGAACGCAATTCAAG 120  
 QY 121 TAATTCCTTGAAGAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180  
 DB 121 TAATTCCTTGAAGAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180  
 QY 181 TAGTCTACTATAATACAGTTTAAATAAGAGTTTATATGAAATTAATTTAATTTGTTG 240  
 DB 181 TAGTCTACTATAATACAGTTTAAATAAGAGTTTATATGAAATTAATTTAATTTGTTG 240  
 QY 241 TTTGCTAGCCTTTGTTGTTTCTTACCTTTGCTCAAGCTGATGTTCTAAAGCGCTT 300  
 DB 241 TTTGCTAGCCTTTGTTGTTTCTTACCTTTGCTCAAGCTGATGTTCTAAAGCGCTT 300  
 QY 301 TTAATTTGGAGCGGAGAGAAACCTTTAGCTTATGAAGTAAAGTAAAGTCTATTG 360  
 DB 301 TTAATTTGGAGCGGAGAGAAACCTTTAGCTTATGAAGTAAAGTAAAGTCTATTG 360  
 QY 361 TGCCATTTTATGAACCTTTTATAGGTTTGAATAGGTTCTTCTCTCAAGAGATA 420  
 DB 361 TGCCATTTTATGAACCTTTTATAGGTTTGAATAGGTTCTTCTCTCAAGAGATA 420  
 QY 421 TTTCTGGGGGTTTCTTCTTCTTCTGATTTGATGAGTTGGTATAGGTTAATTAAGTAA 480  
 DB 421 TTTCTGGGGGTTTCTTCTTCTTCTGATTTGATGAGTTGGTATAGGTTAATTAAGTAA 480  
 QY 481 GAGCTTATTTAGATTCGAAGCTCTTGATAGAAATGCTCAAAACCGCTTTAAGTGA 540  
 DB 481 GAGCTTATTTAGATTCGAAGCTCTTGATAGAAATGCTCAAAACCGCTTTAAGTGA 540  
 QY 541 CTGGGGTAAGGGAATGATGTTGGCAGGTGACATTAATGAGCTGTGACAAAGATTGACAG 600  
 DB 541 CTGGGGTAAGGGAATGATGTTGGCAGGTGACATTAATGAGCTGTGACAAAGATTGACAG 600  
 QY 601 AATATATATTCGTTTACATTTGCTAATAGTTAATAGGAAATGAAATATACCTTA 660  
 DB 601 AATATATATTCGTTTACATTTGCTAATAGTTAATAGGAAATGAAATATACCTTA 660

DB 601 AATATATATTCGTTTACATTTGCTAATAGTTAATAGGAAATGAAATATACCTTA 660  
 QY 661 ATATAGCTTTTGAAGAGGTTTGACCTAGTTTGAATATATATATGAGCCAACTAGCGCTC 720  
 DB 661 ATATAGCTTTTGAAGAGGTTTGACCTAGTTTGAATATATATATGAGCCAACTAGCGCTC 720  
 QY 721 TTGGGTTGAACATCTTTCAAAAAAGTTATTAATTTATTTATTTATTTAAATGAGTGC 780  
 DB 721 TTGGGTTGAACATCTTTCAAAAAAGTTATTAATTTATTTATTTATTTAAATGAGTGC 780  
 QY 781 ATAGCAATTTTGTATGATTTGCTCATTTGTAATTTGAAATATAGAGCTTTGTTATAT 840  
 DB 781 ATAGCAATTTTGTATGATTTGCTCATTTGTAATTTGAAATATAGAGCTTTGTTATAT 840  
 QY 841 TTAATTTTATTTCTGCTGTA 862  
 DB 841 TTAATTTTATTTCTGCTGTA 862  
 RESULT 2  
 ID AAX30100 standard; DNA; 749 BP.  
 AC AAX30100;  
 XX  
 DT 17-JUN-1999 (first entry)  
 XX  
 DE *Borrelia garinii* IP90 protein encoding cDNA.  
 XX  
 KW *Borrelia*; p13 antigen; spirochaete; vaccine; infection; diagnosis; ss.  
 XX  
 OS *Borrelia garinii*.  
 XX  
 FT Key location/Qualifiers  
 FT CDS 192..725  
 FT /\*tag= a  
 XX  
 FN W09912960-A2.  
 PD 18-MAR-1999.  
 XX  
 PE 04-SEP-1998; 98WO-IB01424.  
 XX  
 PR 16-SEP-1997; 97US-0059036.  
 PR 10-SEP-1997; 97DK-0001041.  
 XX  
 PA (SYMB-) SYMBICOM AB.  
 XX  
 PI Bergstroem S;  
 P1  
 XX  
 DR WPI: 1999-215027/18.  
 DR P-PSDB: AAY04280.  
 PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated  
 PT protein p13  
 PS Claim 1; Page 113-114; 118pp; English.  
 XX  
 CC The present sequence encodes a *Borrelia garinii* IP90 protein. The  
 CC present invention describes an isolated nucleic acid (I) that: (1)  
 CC encodes a polypeptide fragment (II) immunologically reactive with  
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of  
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from  
 CC other spirochaetes; and/or (II) hybridizes under stringent conditions  
 CC to specific nucleic acid sequences but not to genomic DNA from most  
 CC other spirochaetes. (I), (II) and transformed cells are useful in  
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.  
 CC (I), (II) and antibodies raised against (II) are used to diagnose such  
 CC infections, in standard immunoassays or amplification/hybridization  
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein  
 CC is involved in virulence and is highly conserved within *B. burgdorferi*  
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those  
 CC responsible for relapsing fever or avian borreliosis). It should provide



**THIS PAGE BLANK (USPTO)**









```

QY 338 AACTAGTAAAGAACCTTATGTCATTTTATGAAACCTTTTATGAGGTTTGGAT 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4435 TATTGAAGTGAATATATTTATATAGAAATATATTTTATGATTTTAT 4494
QY 398 AGGTTCTTTGCTCAAGAGATATCTTGGGGTTTCTTATCTGATTTGATGACGT 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4495 GAATGTAATATATTTTATTTATTTATTTTATTTATTTATTTATTTAT 4554
QY 458 TGGATAGGTTATTTACTTACAGAGCTTATTTAGATTCACAGCTCTGATTAAGAC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4555 TTTTATGCTTGAAGTTTACAAATATATATATATTTTATATATATATATAGG 4614
QY 518 TCCAAAAGCCCTTTTAACTGAGCTTGGGTAAGGAATGATGTGCGAGTGTAC 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4615 AAGATTTATAGTATTTTATTTTATTTTATTTATTTATTTATTTATTTATTT 4674
QY 578 TATGGCTGTGACAGATGACAGAAATTTATTCCTTACATTTGCTAAATGTTATTA 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4675 AATTTATATATATATATATATATATATATATATATATATATATATATAT 4734
QY 638 TAGGAACCTGAAATATACCTTATATATATATATATATATATATATATATAT 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4735 TATTAATATATATATATATATATATATATATATATATATATATATATAT 4794
QY 698 TAAATATGGCCAGCTAGCTGCTTGGCTTGAACCTATCTTCAAAAAAGTTATTA 757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4795 TATAAATATATATATATATATATATATATATATATATATATATATATAT 4854
QY 758 TTATTTTATTTTAAATAGATGATAGCAATTTTGTATGATGCTATGTTATTA 817
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4855 ATTAATGCTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4914
QY 818 AAATTAGACTTTTGTATTTATTTATTTATTTATTTCTCTCTAA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4915 TTTTAACTTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTAT 4959

RESULT 12
LOCUS AX251052 6012 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 20 from Patent WO0168912.
ACCESSION AX251052
VERSION AX251052.1 GI:15984475
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 6012)
AUTHORS Olek, A., Piepdrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and
oncogenes
JOURNAL Patent: WO 0168912-A 20 20-SEP-2001;
EpiGenomics AG (DE)
FEATURES
source
1. 6012
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1842 a 43 c 1153 g 2973 t 1 others
ORIGIN
Query Match 7.8%; Score 67; DB 6; Length 6012;
Best Local Similarity 44.4%; Pred. No. 0.32;
Matches 364; Conservative 0; Mismatches 445; Indels 10; Gaps 2;

```

```

QY 171 ATTTCATTAATATGTTACTATATATACAGTTTATATAAGAGTTTATGA-----A 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4093 ATTTATGTTTATATATTTAGCTTATATGATGTAATAGTGCGTTTATGCGTTTGA 4152
QY 224 TAAATTTTAAATGTTGTTTGTGCTAGACCTTTTGTTGTTTCTTGTGCTCAAGCTGA 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4153 TATTTGTTGTTTGTGTTTGTGATGATTTTATTTTATTTTATTTTATTTTAT 4212
QY 284 TGATTCATAAGCCCTTTTATTTTGGAGCGGAGAAAACCTTTAGCTTATGAACATAG 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4213 TATGAGTGTGTTAGTGTTTTATAGCTATAGTGTATGTTGCTGATGAATTTGATTT 4272
QY 344 TAAGAAATCCATTTTATGCTTATTTTATGAACTTTTATTTTATGAGTTTGAATAGCTTC 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4273 TAGGAGACAGAAATGCGTGTGTTTATTTTGTAGTTTATAGGATGTTTATAGGGGA 4332
QY 404 TTTTGCTCAGGAGATATCTTGGGGCTTTCTTATTTCTGATTTGATGCAAGTTGGAT 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4333 TTTTGTGGGGGTTTAAATTTTATTTTATGTTTGTGTTTATTTTATTTTAT 4392
QY 464 AGGTTATTTACTTACAGAGCTTATTTATATATCAAGCTCTGATTAAGATGCTCAAA 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4393 TTAGATTTTATTTTATTTTATTTTATAGATGTTTAAATTTTATTTTATTTA 4452
QY 524 AGCCGCTTTAATGAGCTTGGGTAAGGAATGATGTGCAAGCTGATTAAGCC 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4453 TTTATAGATTTTATAGGAGGAGGTAATGTTATAGCTTTTGTGTTTATAGTAT 4512
QY 584 TGTGCAAGATGTGACAGAAATTTATTCGTTTACATTTGCTAATTTAATTTAATAG 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4513 ---GATTTATTTTATTTAGTGTATTAATAGTTTAAATTTTATTTTATTTTAT 4569
QY 644 ACTGAAATATAGCCCTTAATATATAGCTTTTGAAGGTTTGAAGCTTGTGATTAATAT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4570 TGTGATTTTATTTGTTATATATATATATATATTTGTTTAAATATTTAAGAGTTT 4629
QY 704 GGCCAGCTAGCCCTCTTGGGTTGAACATCTTTCAAAAAAGTTATTTATTT 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4630 TAGGAAGTTTAAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTAT 4689
QY 764 TATATTAATATAGATGATAGCAATTTGTATGATGTGCTATGTAATTTGAATTTA 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4690 TGTGTTGTTATTTAGCTTTTAAAGTTGTTTATATTTTGGGTATTTTATAGTAA 4749
QY 824 GAGCTTTGTTATTTATTTATTTTATTTTATTTTCTCTCTAA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4750 TTATTTTGTATTAATTTATTTGATTTAGTTATTTTATTTTATTTTATTT 4788

RESULT 13
LOCUS AX344183 6012 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 30 from Patent WO0200926.
ACCESSION AX344183
VERSION AX344183.1 GI:18492071
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Olek, A., Piepdrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with signal transduction
JOURNAL Patent: WO 0200926-A 30 03-JAN-2002;
EpiGenomics AG (DE)
FEATURES
source
1. 6012
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1842 a 43 c 1153 g 2973 t 1 others
ORIGIN

```

```

Query Match      7.88: Score 67.2; DB 6; Length 349980;
Best Local Similarity 47.08; Pred. No. 0.092;
Matches 207; Conservative 0; Mismatches 233; Indels 0; Gaps 0;

1 GATTTTCATTGGATCCAGAAATTTGTAGATTTTCGACAAATTAAGCATTTATTAAAG 60
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY	98	TCGTTGGGAAGCAATTTTCAGTAATCTTTGAAAAAGTTAAATTTAATAGCTTTAA	157
	4203	TTTTTTTGGAAATATATGATTATATTTTGTGATATAAAATATATTAATAATTTTGTGTTA	4262
QY	158	AAACCTTTTAAATTTTCATTAATATGTACTATATATCCAGTTTATATAAGAGCTTTT	217
Db	4263	TGCTTTTAAATTTTGTGTAATTTATTTTTAAATAATGTTTGCTATATATATATTT	4322
QY	218	TATGATTAATTTTAAATGTGTGTTTGCATGACCTTTTGCTTTTCTACCTTGTCTCA	277
Db	4323	ATTAATTTAAATTTTATTTTATAGATATTTGAATTTAAATTTTGTGATTTAAATTTTGAATTT	4382
QY	278	AGCGATGATCTTAAAGCGCTTTTAATTTGGAGCGGAGAAAAACTTTAGCTTATGA	337
	4383	AAAAATATTTATATATAAAAGTTTAAATTTATTTGAGAGA-----TGTTTTTTTAA	4434









```
/note="hypothetical protein; identified by Glimmer;  
putative"
```

Query Match	11.0%;	Score 94.4;	DB 1;	Length 29766;
Best Local Similarity	49.6%;	Pred. No. 1.7e-05;		
Matches 331;	Conservative 0;	Mismatches 326;	Indels 11;	Gaps 3

QY	38	ACAATTAAGCATTTATTAAGAAATTCGAATTCGTAATTTTATGTCGAATCAAGAAGC	97
Db	1921	AAAAATACATAGCATTAATTTCAAAATTTTAAAGAGTTTATGAAACCTTTACTTTAAATTC	1980
QY	98	TCATTTGGGAAGCGAATTTCAAGTAATACCTTTGAAAAAGTTAAATTTAAATAGTTTAA	157
Db	1981	TGCTTAGAAGCAGCAAGAAATTTCAACAACATATCTTTAGAAAAAGATTCACCTCATAT	2040
QY	158	AAACCTTTTAAATTTTCATTAATATGTTACTATATACAGTTTAAATAAAG--AGGTT	215
Db	2041	AAAATCTAAAAATGTGTAAACATTTGTATTCAAAACTTAATTAATTTAGAGGTAAATTT	2100
QY	216	TTTATGAATTAATTTTATTTATTTGTTGTTTGGACGCTTTTGTTGTTTTCAGCTTCCT	275
Db	2101	AATATGAAAAAAATTTTTCACATTAATATTAATTTTATTTAAACAATGCAATCTTTGCA	2160
QY	276	CAAGCTGATGATCTTAAAGCGCTTTTAATTTGGAGCGGAGAAAACTTTTAGCTTAT	335
Db	2161	CAGGTAATACCTTGAAAAAGGTGTGGAGCTGAGATATCGCGACTGTAA---TGAATAT	2216
QY	336	GAACCTAGTAAAGAAAGATCTATTGTGCACTTTTATTTGAACCTTTTATTTAGAGTTTGA	395
Db	2217	GAAGCGAAAAAGCACACATACACTAGCACACTTTGTAATATCTTTTGTCTTTAGGG	2276
QY	396	ATAGCTTTCTTGTCTCAAGAGATATCTTGCGGGGTTTCTTAATCTTGATTTTGATGCA	455
Db	2277	ATAGATATCTTTGTTCAGAGGAGATATATTTGGTGTGTGACACTCTTGCGCTCTCAAGTG	2336
QY	456	GTTGGTATAGGTTTATATCTTACAGAGACTTATTAGATATCAAAAGCTCTGATTAAGAT	515
Db	2337	CTTGGAAAG-----AATCTTATTAATGCGCTGACACATGATCTGAGGAGTATTTGGAAGC	2391
QY	516	GCTCCAAAAGCCGCTTTTAAGTGAAGCTTGGGGGTAAGGAATGATGTTGGCAGGTGCAGTT	575
Db	2392	GTTACAGAAAACACACAGCAAGTATTAATACCGAGAGATATTATTCAGAAATAGGAAGGCTC	2451
QY	576	ACTATNGCTGTGACAGAATTTGACAGAATTTTATATCCGTTTACATTTGCTAATAGTTAT	635
Db	2452	ACAATGTAGACATCTACATATACGTGAATTTATATTCATTTAAATTTGGCAATAGATAC	2511
QY	636	AATAGGAACATGAAAAATAGCCTTAATTTACTTTTGAGAGGTTTGGACCTAGTTTGTAT	695
Db	2512	AACGAGACCTTTAAAAAAGGCTGGCATTGTGCACTTGGGGGGTTGAACCCCATTTTTCAC	2571
QY	696	ATTAATAT 703	
Db	2572	ATTGGAAT 2579	

RESULT	7
AEO01584	
LOCUS	AEO01584 52971 bp DNA linear BCT 21-JUN-2002
DEFINITION	Borrelia burgdorferi plasmid lp56, complete plasmid sequence.
ACCESSION	AEO01584
VERSION	AEO01584.1 GI:6382393
KEYWORDS	.
SOURCE	Lyme disease spirochete.
ORGANISM	Borrelia burgdorferi. Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group. 1 (bases 1 to 52971) Casjens,S., Palmer,N., van Vugt,R., Huang,W.M., Stevenson,B., Rosa,P., Latifga.R., Sutton,G., Peterson,J., Dodson,R.J., Haft,D., Hickey,E., Ghin,M., White,O. and Fraser,C.M. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease
REFERENCE	
AUTHORS	
TITLE	

```

JOURNAL      Spirochete Borrelia burgdorferi
ENTRY        20138354
PUBMED      10672174
2 (bases 1 to 52971)
FRASER,C.M., Casjens,S., Huang,M., Sutton,G.G., Iachigra,R.,
REFERENCE    Fraser,C.M., Dodson,R., Hickey,E.K., Gwinn,M., Peterson,J.,
AUTHORS      van Vugt,R., Palmer,N., Haft,D., Rosa,P. and Stevenson,B.
TITLE        Direct Submission
JOURNAL      Submitted (11-NOV-1999) The Institute for Genomic Research, 9712
SOURCE       Medical Center Dr. Rockville, MD 20850, USA
CURS         Location/Qualifiers
1..52971
/organism="Borrelia burgdorferi"
/db_xref="taxon:139"
/plasmid="lp56"
279..548
/gene="BBQ01"
279..548
/gene="BBQ01"
/gene="BBQ01"
/Note="similar to GB:003641 PID:458212 percent identity:
67.53; identified by sequence similarity: putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAFP07704.1"
/db_xref="GI:6382394"
/translation="MTKRIALIVARFKKDAETICGVRNKTQENYKLIYFLCLLF
GLYGVHREPVKIGIGLLYLPFGFLYGVGLIDLRITTNKFCN"
710..802
/gene="BBQ02"
710..802
/gene="BBQ02"
/Note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAFP07738.1"
/db_xref="GI:6382428"
/translation="MRKKGICGLVIEKSLVGLMFKIVR"
856..1407
/gene="BBQ03"
856..1407
/gene="BBQ03"
/Note="similar to GB:L31615 GB:L31418 PID:520783
PID:551742 percent identity: 97.81; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=1
/product="outer membrane protein, putative"
/protein_id="AAFP07705.1"
/db_xref="GI:6382395"
/translation="WRILVGVFIITALLALCCYCPDNOEQAVQTFEENSISIDGSDSE
IYTEGFSSSLKIVASEHRLVKEIKKVTLLSLDPNRPVPRVDYNEEYFNKFFLDLSS
EESKDLIKFIVKNEQNNKKESEYVWYISLSPDISKESGSESPEDYRMPRP
PAYOOALKRKREIERNTNGKIL"
1404..2268
/gene="BBQ04"
/Note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
GI:520778 percent identity: 93.50; identified by sequence
similarity: putative;protein p23, authentic frameshift"
2744..3496
/gene="BBQ05"
2744..3496
/gene="BBQ05"
/Note="similar to GP:1553115 percent identity: 53.30;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="antigen, p35, putative"
/protein_id="AAFP07706.1"

```

gene  
CDS  
/product="similar to PIR:A64689 percent identity: 35.78;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC66053.1"  
/db\_xref="GI:2690010"  
/translation="MEINLOSGLNNKNNKNIIFISCSLVSPDPNREPTYSKRN  
GVIRPGYKIMKHITLTONNALITTRLSKDRSHAFVSKISLIPGAYEPL  
IOEDSEPERVKKENFKNRHPLDVYNNKFTPEELGITYALYSNIYDRVEEL  
QIDPKIIFVNSDIFVTLGLTDLINSHVAVPTININIGCFSLDETLQNSI  
IEKVFYEETNELYNOTSRFINVSKEVYNTIGSQWTLSTYLRKGRKEMPSVEH  
LEKVIKTIHYTGIOKK"  
2104..2495  
/gene="BBG03"  
/note="This region contains an authentic frame shift and  
is not the result of a sequencing artifact.; similar to  
GP:2239281 percent identity: 63.64; identified by sequence  
similarity; putative; conserved hypothetical protein,  
authentic frameshift"  
complement(2750..2857)  
/gene="BBG04"  
complement(2750..2857)  
/gene="BBG04"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BBG04"  
/protein\_id="AAC66074.1"  
/db\_xref="GI:2690031"  
/translation="MKQFAISIFDLEIAVHIKTLINLNIYNGV"  
complement(2891..4056)  
/gene="BBG05"  
/note="This region contains an authentic frame shift and  
is not the result of a sequencing artifact.; similar to  
GP:2228714 percent identity: 99.12; identified by sequence  
similarity; putative; transposase-like protein, authentic  
frameshift"  
4208..5368  
/gene="BBG06"  
4208..5368  
/gene="BBG06"  
/note="similar to GB:U03641 PID:458216 percent identity:  
66.76; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC66054.1"  
/db\_xref="GI:2690011"  
/translation="MKVFTFLKLCIIYINPISSTMINNSKKNCHNKLKLIY  
LSTLAYNSKYNKTKQNIILKCPENKRRGQPTTTRQNTYIKREKPYTNY  
YKHLGINGETIYKLNYSKRECYLKIQHREKDYFKARVDYLDKRNKGSVD  
LVECLNNNNIKERKLIQIEKYQVIRFKCNLSLKEPLINDINDDELKIL  
KIRRIEINLTKNTLVNSVFEKQNKRLKLSNTQOOLKNGNPEOLEISOKI  
YENKTKHPHFIENHKYKDLSHIKRKLKESIERKENONKOCENNRNTNIFILIQLK  
KEKVIDKPLIKIYLNKNKLEYNKVNNNYVELLEIKKKKLOLKEVY"  
5378..5950  
/gene="BBG07"  
5378..5950  
/gene="BBG07"  
/note="similar to PID:1143398 PID:1174342 PID:1174350  
PID:2182755 percent identity: 57.84; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11

gene  
CDS  
/product="conserved hypothetical protein"  
/protein\_id="AAC66055.1"  
/db\_xref="GI:2690012"  
/translation="MONMAKSIQLVKPIVCSNKKDLFIKIEKDNKTIYTRIMDI  
YKRLNKKKKKKYRISLGLFNOSKVEEENFELRADPGLIYGYKRIKIFVKQ  
VCGLEKSYSLSKTYVMERPKGSIFPCYSLRILKKEVNNTYNNALSMFTLER  
OVIDFDKRTIPQKPFILKIEKSLKHL"  
5911..6678  
/gene="BBG08"  
5911..6678  
/gene="BBG08"  
/note="similar to PID:1174351 percent identity: 39.26;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="stage 0 sporulation protein J (spooJ)"  
/protein\_id="AAC66056.1"  
/db\_xref="GI:2690013"  
/translation="MDRKLAKKHIIAASIKGVKSTSSISTITLSTKNTVLLY  
DLQONAVTSYFTQDHPRLMLNIYNSYSLIKHKTFKDVVISKNDLPISYEL  
AKFSKGNQFRELNRNAVNYLDEYDVIIDPPSSSELNMLVADVILIPLE  
RMAVENPLLIQIKLELNEMGKEAKIHFASKVEIGRTSTELMSLKEKLNKE  
IGEVHSEALKIHDVAIGPENENYKEYLRILEKI"  
6737..7288  
/gene="BBG09"  
6737..7288  
/gene="BBG09"  
/note="similar to PID:1143400 percent identity: 43.68;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC66057.1"  
/db\_xref="GI:2690014"  
/translation="MWKRVYIINDRIYRNATYINTERKREYELKLNELNRIEDD  
IRKNTMKILLETIRNKTYLIDGYKKREDFIPDKIARQAVYIKIAXLIFGKLE  
EIIDENGIDKTLLENLAKDKINKANLITPLRPLEQECDDYKNNPKRANLLED  
FYORKNEOLIKLEEKOKYS"  
complement(7483..10779)  
/gene="BBG10"  
complement(7483..10779)  
/gene="BBG10"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BBG10"  
/protein\_id="AAC66075.1"  
/db\_xref="GI:2690032"  
/translation="MKLDEIILPLSMISISNOQKDAIETIKKIAEOKFSSLDLKS  
LEKSAKSGSNLEKALNKSQNAIKNFSLANSVNSKAGSKSIGVLSKVGGLG  
SVKLNANKTGAFNOMLAAPIIYAVKAIQAGSTTIGIDGAMDADEFEENSTF  
SDMIGNEDIGKSLAESMRAGFDETLFTDAIANATKMTLSGATVASEVERIRMGGA  
AGCSSECELEKLAEYVRSSESNQVNLDELVALRAGVDTPIILAEGLAEALYKKA  
SDGKIGFALNKALSKATSEGGKFTGNTAKAKTLQAOQOOTAKSEKFLDIGKALE  
PMIIGFEVKQFLWALVWPLKVTATYITTKGELVAVLAGWVGKMLERAIT  
TFPVLYEGIVYIGKLELATYVSSGLDFKGFEBPIIKWOKLIDMISKATYKLN  
NLVPEWKAKEKNDISGESEPRDKFDENATNINKAAEDYOKLODIENRQDIYN  
KTKARADOLANLEKTIENKOKKDEKYSLEDTLDENKILVGEKSVNEPNSNY  
DPVNETOULLKESREREITIKTLPHTQVYALOKLNEITIEKKAPFEYKGSPELT  
NESNQVVALEKOVNEYEKALDLSFVDAQALOKETTDLEMETALLPAKERSAEK  
KMAKSIQAMYKRPVDEHRSQPKRINLETNRNTIKOYEAQDPTTSLVDSMLDGLNVE  
NAFMKIDIAKFLNMDTGESIEEFHNLINGDVNMBEGLMPTQOMESMTGKLTAA  
GAVRPGAEVAELNGLDPEVWGLKGOEKRIKAIKIKDEDELEEKSEYELK  
LEDREDEIKMKREKLSLDEYVTEIFELQAOAKGKQISGEEOKRLHDVOTEYKTK  
KDIEFTQITKTEBAKLEVERKNTSKIEPRIRIQAQAVDVYNAEDMNLGKPDRLK  
TQKILDEILKRIAKVSAGEITKLHAGGARFVSNKPTYPNSGWSSEFGQPELVK  
ITPAIDENLRLKLEAKIIAETITKLOKQSSVYNNFYNGVDLAEKLVRLKSK  
EHIMGFRAAE"  
complement(10776..11015)  
/gene="BBG11"  
complement(10776..11015)  
/gene="BBG11"

```

gene      KNEHNSILFTLINEAIGIKNSAIAIHVCENID"
           complement(4623..4721)
           /gene="BB111"
           complement(4623..4721)
CDS       /gene="BB111"
           /note="hypothetical protein; identified by Glimmer;
           putative"
           /codon_start=1
           /product="B. burgdorferi predicted coding region BB111"
           /protein_id="AAC66199.1"
           /db_xref="GI:2690096"
           /translation="MSKLGYSFNVVPEPKIQIKLIYNKGSVLE"
           5128..5346
           /gene="BB112"
           5128..5346
           /gene="BB112"
           /note="hypothetical protein; identified by Glimmer;
           putative"
           /codon_start=1
           /product="B. burgdorferi predicted coding region BB112"
           /protein_id="AAC66200.1"
           /db_xref="GI:2690097"
           /translation="MIKEPKNHISKIHIEKEYEDVTIKWDREYSIFRKLHGKNTLED
           MLEYYOKEENOKIKEFANKFIKKRPKI"
           5609..5707
           /gene="BB113"
           5609..5707
           /gene="BB113"
CDS       /gene="BB113"

```

Query Match 11.6%; Score 100; DB 1; Length 27323;

Best Local Similarity 50.5%; Pred. No. 2.6e-06;

Matches 410; Conservative 0; Mismatches 355; Indels 47; Gaps 5;

```

QY 39 CAAATAAGACATTTATAAAGAATGAAATGCTAATTTATGTCAAATCAAGAAGCT 98
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20290 CTAATAGAGACTCTATTGATCATCAAAAAGTATATTTAAATCAAAATTAACCTAT 20231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 99 CATTGGGAAGCAATTTCAAGTAATACCTTTGAAAAAGTAAATTAATTAAT 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20230 TATTACAAACCTCATCTTAACCCAAATTCATGAAATTAATAGTAAATCTTTGG 20171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 158 AACCTTTTAAATTCATTAATATGTTACTTAATACAGCTTTTAATTAAGAGCTTT 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20170 GGGTTTGTATATATGCTATATTAATTTAAATTTTGGAGGTAAATATCA 20111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 218 TATGAATTAATTTAATTTGTTGCTAGCCCTTTGTTTCTAGCTTGTCTCA 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20110 TATGAAAAAATTTTAACATTTGATTAATTTTACCTTAACAATCAAAATCTTGAAC 20051
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 278 AGCTGATGATCTTAAGAAGCCCTTTTAATTTGGAGCGGGAAGAAACCTTTACCTATCA 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20050 ACAAGATTAAGCTTGAAGAAAGCTT-----GGAAGTATTAAGAACATTAAGAAATAA 19997
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 338 AACTAGTAAGAAAGATCTATTGTCATTTTATGAACCTTTTATAGGTTTGAAT 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19996 AAGCGAAAAAGCACTATATAGCACCATTCTTTGAATTTATTTTAACTTTAGGAAT 19937
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 398 AGGTTCTTTGCTCAAGAGATATCTTGGGGGCTTTCTTAATCTTGATTTGATGCAGT 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19936 AGGATCTTTGCTCAAGAGATATATGTTGGTGGCGGCGCTCTTGATCTCAGTTAT 19877
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 458 TGGTATAGGTTAATCTTACAGAGACTTATTTAGATCAAAAGCTTGTATAGAAATGC 517
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19876 AGGAGGAATACCTTTCATAGCTGGAATATCTTGGCCATATGATGATGAACAAAGAGC 19817
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 518 TCCAAAAGCCGCTTTTAAGTGACTGGGCTAAGGAGATGTTGGCAGGTGCAGTTAC 577
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19816 AACAA-----CTGGGCTATATATTAACACATAGAGATAGGCAC 19778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 578 TATGGCTGTGACAAGATGACAGAAATTTATTCGTTTACATTTGCTAATAGTTATA 637
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 19777 GATTATAGCATCCACATACCTCACTTATTTATTCATTTACATTTGCCAATAACACA 19718
QY 638 TAGGAACAGTGAATAATAGCTTAAATATAGCTTTGGAGGGTTTGAGCCTGTTGATAT 697
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19717 TGCAAATCTTAAAAAAGACTCGGCATTTGATATGCGGGGTTTGAACCCATTTGATAT 19658
QY 698 TAATATGGCCCAAGCTAGCCCTCTGGGTTTGAAGTATCTTTCAAAAAAGTTATTAAT 757
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19657 TGAATATAGC-----GGATTCCAACTATCGTTTAAAAAAGATATCAAAAT 19613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 758 TTTATTTA-----TTATTAATATGATATAGCAATTTTGTATGTGATTCCTATTTGTA 813
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19612 AAATATCAAGCTATAAAAATTATGTTACAAAAAAGTACTAGTATATAAATAATAA 19553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 814 TTGAATATAGCTTTTGTATTTATTTATTA 845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19552 ATTTAATTAATGCAATTTTATTAATAATATA 19521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6  
AE000786 29766 bp DNA linear BCT 17-DEC-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
Borrelia burgdorferi.  
Borrelia burgdorferi.  
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia  
burgdorferi group.

1 (bases 1 to 29766)  
Fraser C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.,  
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K.,  
Gwinn M., Dougherty B., Tomb J.F., Fleischmann R.D.,  
Richardson D., Peterson J., Kierlavage A.R., Quackenbush J.,  
Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D.,  
Gocayne J.D., Weidman J., Uterback T., Wathey L., McDonald L.,  
Artlich P., Bowman C., Garland S., Fujii C., Cotton M.D., Horst K.,  
Roberts K., Hatch B., Smith H.O. and Venter J.C.  
Direct Submission  
Submitted (08-DEC-1997) The Institute for Genomic Research, 9712  
Medical Center Dr. Rockville, MD 20850, USA

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

2 (bases 1 to 29766)  
Fraser C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K.,  
Gwinn M., Dougherty B., Tomb J.F., Fleischmann R.D.,  
Richardson D., Peterson J., Kierlavage A.R., Quackenbush J.,  
Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D.,  
Gocayne J.D., Weidman J., Uterback T., Wathey L., McDonald L.,  
Artlich P., Bowman C., Garland S., Fujii C., Cotton M.D., Horst K.,  
Roberts K., Hatch B., Smith H.O. and Venter J.C.  
Direct Submission  
Submitted (08-DEC-1997) The Institute for Genomic Research, 9712  
Medical Center Dr. Rockville, MD 20850, USA

FEATURES

source

gene

CDS

1. 29766  
/organism="Borrelia burgdorferi"  
/db\_xref="taxon:139"  
/plasmid="lp28-2"  
116..1009  
/gene="BBG01"  
116..1009  
/gene="BBG01"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BBG01"  
/protein\_id="AAC66073.1"  
/db\_xref="GI:2690030"  
/translation="MKSLFLVTLIMGIMSCNLSKSSNKEQNNNNVKEYSNSVQ  
EDGLNDVLSNDEKRSFTNKGEMWYEDLINIEPIISESRKNANIPNIEIVHOK  
KIKEDDIPSTNERKEADEAIKYLEENILNOSKSELSIREVRLKDEVALINSDFD

Gwin, M., Dougherty, B., Tomb, J.-F., Fleischmann, R.D.,  
 Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J.,  
 Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D.,  
 Goeyne, V.D., Weidman, J., Utterback, T., Walthey, L., McDonald, L.,  
 Artach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K.,  
 Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.  
 Direct Submission  
 Submitted (08-DEC-1997) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 Location/Qualifiers  
 1. 27323  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 /plasmid="lp28-4"  
 174. 608  
 /gene="BB101"  
 174. 608  
 /gene="BB101"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB101"  
 /protein\_id="AAC6211.1"  
 /db\_xref="GI:2690108"  
 /translation="MENQYFNFIFPKIIFILYIMKSVKKEQEKIDHACRIIL  
 LTFITFENILLENYSOKTLKYNENKLNPPNSITPMKKYLKQKEKIKIIAK  
 YEKNDOSLIYKKLVLTLEKTLKLELFYKELKQFIQKNTT"  
 736. 1104  
 /gene="BB102"  
 736. 1104  
 /gene="BB102"  
 /note="similar to GP:1655791 percent identity: 66.67;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAC6219.1"  
 /db\_xref="GI:2690086"  
 /translation="MDGVINDTIVARMKKQIKFNKNTLITVKTLDHKNKELYSANK  
 TYNVLIQNNNEALAKTYQLRVNKKTLLELLEKPKVILKRPYKKNESFGLY  
 TLLSPLEYCCYKTYNLHPNT"  
 complement(1847. 1972)  
 /gene="BB103"  
 complement(1847. 1972)  
 /gene="BB103"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB103"  
 /protein\_id="AAC6212.1"  
 /db\_xref="GI:2690109"  
 /translation="WVYVLDPNLSLYFGGALCYWDFYVWVWISLSEIVSFL"  
 complement(2124. 2219)  
 /gene="BB104"  
 complement(2124. 2219)  
 /gene="BB104"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB104"  
 /protein\_id="AAC6213.1"  
 /db\_xref="GI:2690110"  
 /translation="WGTIQKFLIKHGLMRAKPIGESKYFIY"  
 complement(2188. 2310)  
 /gene="BB105"  
 complement(2188. 2310)  
 /gene="BB105"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB105"  
 /protein\_id="AAC6214.1"  
 /db\_xref="GI:2690111"  
 /translation="MPYLVNNSNVOKETILKALKFFODRKSIGNNPKIFN"  
 2536. 3351  
 /gene="BB106"  
 2536. 3351  
 /gene="BB106"  
 /note="similar to GB:026562 SP:P24247 GB:M83735 PID:473817  
 PID:457107 percent identity: 33.62; identified by sequence  
 similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="pfs protein (pfs)"  
 /protein\_id="AAC6219.1"  
 /db\_xref="GI:2690087"  
 /translation="MKIKNINILIFLSIVLNSMGNSKINTPENSOSNIIISAT  
 KAEIEINKIIONKKYISIEHRRKKIAIGKLMDSNITITATGVKINTALMTSYII  
 SKYKISHINAGVASGYSIDKKNFKIGDVISETYSYEDLHFGYEIGHVEHPK  
 KEKANTALIRKTSKIKINNTSYMGILITGDFIDHOFORIPPEFEFAIADMSAA  
 MAQVAYGKPIPIIRIGISDVNNENNNDYKRLKASSASSAKIVENTILKM"  
 complement(3343. 3441)  
 /gene="BB107"  
 complement(3343. 3441)  
 /gene="BB107"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB107"  
 /protein\_id="AAC6215.1"  
 /db\_xref="GI:2690112"  
 /translation="WQLKGAIIKLEDEYKIKIQLNLIKILH"  
 3576. 3735  
 /gene="BB108"  
 3576. 3735  
 /gene="BB108"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB108"  
 /protein\_id="AAC6216.1"  
 /db\_xref="GI:2690093"  
 /translation="MIIISFPDYNPSASNIFIKITTPKCLAKKLLPLQKQFNNLPS  
 FYKNIKYSRSLNICK"  
 3745. 3882  
 /gene="BB109"  
 3745. 3882  
 /gene="BB109"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB109"  
 /protein\_id="AAC6217.1"  
 /db\_xref="GI:2690094"  
 /translation="WASBAKIDAESKNEVNSQKSKNSKNIEVKDTGLVKAIRKS  
 L"  
 3911. 4315  
 /gene="BB110"  
 3911. 4315  
 /gene="BB110"  
 /note="hypothetical protein; identified by glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB110"  
 /protein\_id="AAC6218.1"  
 /db\_xref="GI:2690095"  
 /translation="WGRAMTPHQQLKYNLEGLKMINLDELKIAIDGNDNNISKY  
 NDLKKVDFNPNNSLIKYVLYLKNSSNEDKIEAKICKITLMHNVEYFESVDELDKDN

QY	575	TACATATGCGCTGACACAAATGACGAATATATATTCGCTTACATTTGCAATAGTTA	634
Db	481	TACTATGCGCTGACACAAATTAACGAATATATTTCTTCATATTTACATTTGCTAATAGTTA	540
QY	635	TAATAGAAACTGAAAAATAGACCTTAATATAGCTTTTGAGAGGTTTGACCTAGTTTGA	694
Db	541	TAATAGGAAGCTAAAAAATAGCCTTAATATAGCTTTAGAGAGATTTGAACCTAGTTTGA	600
QY	695	TATTAATATATGGCCAGCAGCAGCTCTTGCGTTTAACTATCTTTCAAAAAAGTTATTA	754
Db	601	TGTTGCAATATGGCCAGCAGCTCAGTCTTGCGTTGAACTGCTTTCAAAAAAGCTATTA	660
QY	755	ATTATTTATTTATTTAATAATGATGATAGCAATTTTGATTTGATGATCTC-ATTGTAA	813
Db	661	ATTTTA-TTTATTTACAAAAATGCGGATTCGATTCGATTTCTGATTTGCAATTCGTCATTTGTA	719
QY	814	TTGAAAATTAGAGCTTTTGTTTATTTATTTAATTTTA	850
Db	720	TTAAAAAGACTTTTACTTTGTTATTTTATTTCTCTACTA	756
SQUIT_4			
LOCUS	AF085741	678 bp	DNA linear BCT 24-APR-2001
DEFINITION	Borrelia garinii membrane protein p13 (p13) gene, complete cds.		
ACCESSION	AF085741		
VERSION	AF085741.1	GI:4731124	
KEYWORDS			
SOURCE	Borrelia garinii.		
ORGANISM	Borrelia garinii.		
	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia		
	burgdorferi group.		
REFERENCE	1 (bases 1 to 678)		
AUTHORS	Noppa,L., Ostberg,Y., Lavrinovich,M. and Bergstrom,S.		
TITLE	P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains		
JOURNAL	Infect. Immun. 69 (5), 3323-3334 (2001)		
MEDLINE	21189251		
PUBMED	11292755		
REFERENCE	2 (bases 1 to 678)		
AUTHORS	Noppa,L., Ostberg,Y. and Bergstrom,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea		
	University, Umea S-901 87, Sweden		
FEATURES	Location/Qualifiers		
source	1..678		
	/organism="Borrelia garinii"		
	/strain="ip90"		
	/db_xref="taxon:29519"		
	/chromosome="large linear"		
gene	121..654		
	/gene="p13"		
CDS	121..654		
	/gene="p13"		
	/codon_start=1		
	/transl_table=1		
	/product="membrane protein p13"		
	/protein_id="AAD28362.1"		
	/db_xref="GI:4731125"		
	/translation="MNKFLIFLIVFCAFSSEADDSKSPFNIGAGEKELYETNNKDD		
	SLVPELNLFLFGIVTFAOGDILIGSSILIEFDVVGILITGAYLIDKFDNNAKRAKRA		
	DFKTMKCGKMLAGVGMVAATVRLTEIYLPEFANNYNRKLTNSINIALGGEPEPSDIN		
	MGQASALGFGLSFKKSY"		
BASE COUNT	202 a 77 c 140 g 259 t		
ORIGIN			
Query Match	62.9%; Score 542.4; DB 1; Length 678;		
Best Local Similarity	89.6%; Pred. No. 1.4e-71;		
Matches 606; Conservative 0; Mismatches 66; Indels 4; Gaps 2;			
98 TCTATTTGGAAAGCAATTTCAAGTAATCTTTGAAAAAGTTAAATTAATAGTTTAA	157		

Db	1	TCTATTGGGAAGCCAAATTTCCGATGACAAATTTGAAAAAGTTAAATTTAAATAGTTC	60
OY	158	AAACCTTTTAAATTTCAATTAATATGTTACTATATATACCACTTTTAAATAAGAGCTTT	217
Db	61	AAGCTTTTAAATTTCAATTAATATGTTACTATATATACCACTTTTAAATAAGAGAGCTTT	119
OY	218	TATGAATAAATTTTAAATTTGTTGTTGCTGAGCCCTTTGTTGTTTTCAGCTTGCTCA	277
Db	120	TATGAATAAATTTTAAATTTTATTTTGGTAATCTTTTGCTTTTCTAGCTTTGCTCA	179
OY	278	AGCTGATGATTTCTAAAGACGCTTTTAAATTTGGGACGGGAGAAAACTTTAGCTTATGA	337
Db	180	A---GATGATTTCTAAAGACGCTTTTAAATCTGGACGGGAGAAAAATTTTGGTTATGA	236
OY	338	AACTGATGAAGAAGATCTCTATTTGCTCATTTTATTTGAACCTTTTAAAGGTTGGAAAT	397
Db	237	AACTGATGAAGAAGATCTCTCTGACATTTTATTTGAACCTTTTAAAGGTTGGGAT	296
OY	398	AGGTCCTTTTGGTCAAGGAGATATCTTGGGGGCTTTCTTATTTCTGATTTGATGACAGT	457
Db	297	AGGTCCTTTTGGTCAAGGAGATATCTTGGGAGCTTTCTTATTTCTGATTTGATGACGCT	356
OY	458	TGGTATAGGCTTAAATCTTACAGAGCTTATTTAGATATCAAGCTCTTGATAGAAATGC	517
Db	357	TGGTATAGGCTTAAATCTTACAGAGCTTATTTAGATATCAAGAGATTTTGATATATATGC	416
OY	518	TCCAAAAGCCGCTTTTAAAGTGAGCTTGGGGATAGGGAATGATGTTGCGACGCTGACCTTAC	577
Db	417	TAAAAAGCTGATTTTAAAGTGAGCTTGGGGATAGGGAATGATGTTGCGACGCTGCTTAC	476
OY	578	TATGGCTGAGCAACATTTGACGAATATATATTCGCTTACATTTGCTAAATAGTTATAA	637
Db	477	TATGGCTGAGCAACATTTGACGAATATATATTCGCTTACATTTGCTAAATATATATAA	536
OY	638	TAGGAACCTGAAAAATAGCCCTTAATATATAGCTTTTGAGAGGTTTGAGCCTAGTTTGATAT	697
Db	537	CAGGAAGCTGAAAAATAGCTTAAATATATAGCCTTGAGAGGATTTTGAGCCTAGTTTGATAT	596
OY	698	TAAATATGGCCAGCTAGCAGCTCTTGGGTTTGAACTATCTTTCAAAAAAGTTATTAATTT	757
Db	597	TAAATATGGCCAGCTAGCAGCTCTTGGGTTTGAACTATCTTTCAAAAAAGCTATTAATTT	656
OY	758	TTATTTTATTTATTAATAA 773	
Db	657	TTATTTTATTTATTAATAA 672	

RESULT 5

AE000789/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

27323 bp

DNA

linear

BCT 17-DEC-1997

Borrelia burgdorferi plasmid lp28-4, complete plasmid sequence.

AE000789

AE000789.1 GI:2690079

Borrelia burgdorferi.

Borrelia burgdorferi.

Bacteria: Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.

1 (bases 1 to 27323)

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J., Venter, J.C. et al.

Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi

Nature 390 (6660), 580-586 (1997)

98065943

9403685

2 (bases 1 to 27323)

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,

Matches 740; Conservative 0; Mismatches 106; Indels 5; Gaps 3;

```

OY 4 TTTTCATGAGTCCAGAAATTTAGAAATTTGCAACAATTAAGACATTAATAAGAT 63
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6646 TTTTCATGAGTCCAGAAATTTAGAAATTTGCAACAATTAAGACATTAATAAGAT 6587
OY 64 TGAATGCTAATTTATGTCGAATCAATCAAGAGCTAATGGGAAGCAATTTCAAGTAA 123
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6586 TGAATGCTAATTTATGTCGAATCAATCAAGAGCTAATGGGAAGCAATTTCAAGTAA 6527
OY 124 TACTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAAATTCATTAAT 183
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6526 TAAATTTGAAAAAGTTAAATTTAAATTAACCTTTTAAATTCATTAAT 6467
OY 184 GTTCTATATATACAGTTTATATAAAGCTTTTATGAATTAATTTAATGTTGTT 243
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6466 GCTACCATATACAGTTTATATAAAGGCTTTTATGAATTAATTTAATTTGTTT 6407
OY 244 TGCTAGCCTTTTGTGTTTTTCTAGCTTGTCAAGCTGATGATCTTAAA--GCCTT 300
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6406 TGCGAACCTTTTGTGTTTTTCTAGCTTGTCAAGCTGATGATCTTAAAATGTTGCT 6347
OY 301 TTAATTTGGAGCGGAGAAAACTTTTATGCTATGAAACTAGTAAAGATCCTATG 360
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6346 TTGGATGAGTGTGAGAAAACTTTTGTATGAAACTAGCAAGACAGATCCTATG 6287
OY 361 TGCAATTTTATGACCTTTTATAGGCTTGTGAATAGCTTCTTTCCTCAGAGATA 420
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6286 TACCAATTTTATGACCTTTTATAGGCTTGTGAATAGCTTCTTTCCTCAGAGATA 6227
OY 421 TTTCTGGGGGTTTTCTTATTTCTGGAATTTGATGAGTGTATAGGTTAATCTTACAG 480
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6226 TTTCTGGAGGTTCTCTATTTCTTGATTTGATGCGGTTGATAGGCTTATCTTGCGG 6167
OY 481 GAGCTTATTTGATATCAAGCTCTTGATTAAGATGCTCCAAAACCGCTTTAAGTGA 540
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6166 GGGCTTATTTGGAATCAAAAGCGCTTGTATCTACAAAAAGCTTTCAATGGA 6107
OY 541 CTGGGGGGAAGGAGATGTTGGCAGGTGAGTACTATGCTGTCAGCATGATGACAG 600
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6106 CTGGGGGGAAGGAGATGTTGGCAGGTGAGTACTATGCTGTCAGCATGATGACAG 6047
OY 601 AAATTTATTTCCGTTTACATTTGCTAATAGTATTAATAGGAAAACTGAAGCCTTA 660
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6046 AAATTTATTTCCGTTTACATTTGCTAATAGTATTAATAGGAAAACTGAAGCCTTA 5987
OY 661 AATATGCTTTGGAGGCTTTGAGCCTACTTTGATATTAATAGGCCCAAGCTAGCCTC 720
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5986 ATGATGCTTTGGAGGCTTTGAGCCTACTTTGATATTAATAGGCCCAAGCTAGCCTC 5927
OY 721 TTGGGTTGAGCTATCTTCAAAAAAGTATTAATTTTATTTTATTTAAATGAGTGG 780
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5926 TTGGGTTGAGCTATCTTCAAAAAAGTATTAATTTTATTTTAAATGAGTGG 5868
OY 781 ATAGCAATTTGTATGTGATGCTC-ATGTAATTTGAATAATAGAGCTTTGTTTATTA 839
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5867 ATGCAATTTGTATGTGATGCTC-ATGTAATTTGAATAATAGAGCTTTGTTTATTA 5808
OY 840 TTTTATTTTA 850
    ||| | | | | |
Db 5807 TTTTCTACTTA 5797

```

RESULT 3  
 AF085739 781 bp DNA linear BCT 24-APR 2001  
 DEFINITION *Borrelia burgdorferi* membrane protein p13 (p13) gene, complete cds.  
 ACCESSION AF085739  
 VERSION AF085739.1 GI:4731120  
 KEYWORDS  
 SOURCE  
 ORGANISM  
*Borrelia burgdorferi*.  
*Borrelia burgdorferi*.  
 Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*; *Borrelia*

REFERENCE  
 Noppa, L., Ostberg, Y., Lavyrinovicha, M. and Bergstrom, S.  
 p13, an integral membrane protein of *Borrelia burgdorferi*, is  
 C-terminally processed and contains surface-exposed domains  
 Infect. Immun. 69 (5), 3323-3334 (2001)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 21189251  
 11292755  
 REFERENCE  
 2 (bases 1 to 781)  
 Noppa, L., Ostberg, Y. and Bergstrom, S.  
 Direct Submission  
 Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea  
 University, Umea S-901 87, Sweden  
 FEATURES  
 source  
 1. 781  
 /organism="Borrelia burgdorferi"  
 /strain="B31"  
 /db\_xref="taxon:139"  
 /chromosome="large linear"  
 122..661  
 /gene="p13"  
 122..661  
 /gene="p13"  
 /note="similar to *Borrelia burgdorferi* predicted coding  
 region B80034 encoded by Genbank Accession Number  
 A801117"  
 /codon\_start=1  
 /transl\_table=11  
 /product="membrane protein p13"  
 /protein\_id="A028360.1"  
 /db\_xref="GI:4731121"  
 /translation="MNKILFVATFCVPSFPAQNDNRKAGMSAGEKLYETSK  
 QDPIVPLMLPLGFGISFAGODIGLSILGFDVIGLILAGAVLDIKADGCTK  
 KAPQRTWKGQWLAGVYVTAATRLRIEILIPFTANSYNRKLNSVNLGREFPSFD  
 VAMGSSALDFELSFKSY"  
 BASE COUNT 223 a 98 c 153 g 307 t  
 ORIGIN

Query Match 65.2%; Score 561.8; DB 1; Length 781;  
 Best Local Similarity 86.5%; Pred. No. 1.8e-74;  
 Matches 655; Conservative 0; Mismatches 97; Indels 5; Gaps 3;

```

OY 98 TCTATTGGGAAGCAATTTCAAGTAAATCTTGGAAAAAGTTAAATTAATGTTTAA 157
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TCTATTGGGAAGCAATTTCAAGTAAATCTTGGAAAAAGTTAAATTAATGTTTAA 60
OY 158 AAACCTTTTAAATTTTCAATTAATATGTTACTATAATTAACAGTTTAAATAGGTTT 217
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 AAACCTTTTAAATTTTCAATTAATATGTTACTATAATTAACAGTTTAAATAGGTTT 120
OY 218 TATGAATTAATTTTAAATTTGTTGTTTGGTACGCTTTTGTGTTTCTAGCTTGTGCTCA 277
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TATGAATTAATTTTAAATTTTGTGTTTGGCAACCTTTTGTGTTTCTAGCTTGTGCTCA 180
OY 278 AGCTGATGATTTCTAAA--GCCTTTTAAATTTGGAGCGGAGAAAAACTTTTACGTTA 334
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AGCTAATGATTTCTAAAATGTTGGCTTTGGGATGAGTGTGAGAAAAACTTTTGTGTTTA 240
OY 335 TGAACCTAGTAAAGAAAGTCTATTTGCCCATTTTATTTGAACTTTTATAGGTTTGG 394
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TGAACCTAGTAAAGAAAGTCTATTTGCCCATTTTATTTGAACTTTTATAGGTTTGG 300
OY 395 AATAGGTTCTTTGCTCAAGAGATATCTTGGGGGCTTTTCTTATTTCTGATTTGATGC 454
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 AATAGGTTCTTTGCTCAAGAGATATTTCTTGGAGGTTCTTATTTCTGATTTGATGC 360
OY 455 AGTTGGTATAGGTTAATATCTTACAGAGCTTATTTAGATATCAAAAGCTCTGATAAGAA 514
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GGTGTGATAGGCTTATATCTTGGCGGCTTATTTGATATCAAAAGCTCTGATGATAT 420
OY 515 TGCTCCAAAACCGCTTTTAAAGTGTGAGTGTGGGATGAAGGATGATGTTGGAGGTGAGT 574
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 TACTAAAAAAGCTCTTTTCAATGAGCTTGGGATGAAGGAGTATATGTTAGCAGGTGTGT 480

```



```

putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0028"
/protein_id="AAC6428.1"
/db_xref="GI:2687919"
/translation="MKQRYENFKKRLILNLLIFLLACSSSEIFSOGLNOKIKHEY
NILGSSPRGISIVGETLYIAMHLFKKENGKIEKIDLSNYSFEINDIYNISGKTYLL
AONKEEELVEELNKGDMTLKPKKPLKAYFKSVGRGVKAYILAIADKNKREKIFD
LOGSDKTPPOATENDKKEOISNEENLTGNSLKIQMNNNTYTNIDYQOAKREIPITK
TSIRGSSSEVLVMTGCTYNNLDTRFKYVSTNNNTTTFITIDEVGEFSYFARFENAIL
IGSNNGFAEFKMKEGIFALRAPSVEBPAGVNGSOLSKTGILNDIIPVSNNTIYLLTG
GKGLMKLENKRLTKE"
/complement(1776. .2207)
/gene="BB0029"
/complement(1776. .2207)
/gene="BB0029"
/note="similar to SP:PI7868 GB:X03897 PID:40141
PID:1303831 GB:AL009126 percent identity: 32.58;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC6423.1"
/db_xref="GI:2687914"
/translation="MLSKIFVDADSCNFKIILQKFIHLNKSRLIVSNKPFSEIT
ENVALEVVDNDSFILNLADRSIVITRDIFPAKLLDLGVYVMDDEGRIDINNINY
LYFRSKINPNLKIKIKKYYDEDFNKSREKFTTNYSLFPS"
/complement(2197. .2832)
/gene="BB0030"
/complement(2197. .2832)
/note="similar to GB:236269 PID:1049113 percent identity:
30.95; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="signal peptidase I (lepB-1)"
/protein_id="AAC6422.1"
/db_xref="GI:2687913"
/translation="MAPYLTFEORLLRRKORRVPEKVFLETLINFEPTKPYLQIFMI
KSNEMLPITTKNASLFFVATHTTSFFIPLKANDIYLFEDFRLSNPFLTLTKDFEFLN
KIFKASIKVSRILAVQDSVYRGVLNVKNDNFYLVNGLNLSYKRLKDNFMTDE
VICKETLKNEFFLNDMLSVLNDSRIRGPIKNAIVSEFLVKVDYKIVK"
/complement(2834. .3814)
/gene="BB0031"
/complement(2834. .3814)
/note="similar to PID:1652260 percent identity: 35.98;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="signal peptidase I (lepB-2)"
/protein_id="AAC6421.1"
/db_xref="GI:2687912"
/translation="WYLRRLDKFASFLVHIVKLYTRKKRYFCKLRAKRGFLNLF
LFEVAAALFVLVINOYVOAQLKIPSGSMENLQIGDLFVDRKSYGPELPLGKLEKIN
GAPADESDIILFENPEYKSGVPEIDIRLITLMTLSFIDLRDEGPNPVRYKVR
GAPADGKIVRNNNGKAYIKRGEENFILEDSDWDYDNKFKTKIYANEDGIGTYCDFA
MVALSOLNINLSSTPDFSFDVRMIDREFEERLETKYLSAMPYVDVYMEALMRDY
GIYPDGIILPIGDRNDSHDGREFGVINKKVLGRTLIYLFPSRVGFI"
/complement(3885. .5357)
/gene="BB0032"
/complement(3885. .5357)
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0032"
/protein_id="AAC6427.1"
/db_xref="GI:2687918"
/translation="MKEDVENSVELFEVKLPLIGIEPKVYVFLTIIILLISLTSL
IIPKFNPGAYLKNINSNIENWTYIYNKYGIRPPLNKYINATCEVLAKAKRGFTYE

```

```

ORIKIHNKFCNGYSLOINLELVDEPEKIKOKRELISIMWKIKININENTRLIIPFESLIS
SELKPKYIKKLEKDSIPYLNSTEMFENPISIDONSNOENENISLKTNF
DLNRAIFPELLENLDKDIKITRNAPWKTLAKTDIANDINOLISKEINIKLPKFK
INSNKIEKIÖNELNSLDSKNSLSTYNNKFEILQEOENVTRYEIOEDLKEPKKALN
KNEMLIKEDLVENLKNENÖIGLEAATIGISYFAITAMYSKRLPTGFEARLPIIS
ÖEWELÖKPEPNKPLNINELISKRVGMFMOMONSNEAIFENKKNFYSENSNFYSLI
TEIRFYSOÖNNLNASTKASFLKMSSPNIGFRILVSE"
5369. .5821
/gene="BB0033"
/complement(5369. .5821)
/note="similar to GP:2073503 percent identity: 47.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="small protein (smpB)"
/protein_id="AAC6420.1"
/db_xref="GI:2687911"
/translation="MNTNNLLLENKAKFNFFIEKISGIVLKGTEVKSIAKKLSTF
NNSFASIKKEELMELNLSVSKYKEGNFNHDELRRKRLIKRELÖRLKFKREKGYT
LIPISFYLKSITIKYEVGICGKKLYDKRELKOKSIKKDLISREIKYK"
/complement(5892. .6431)
/gene="BB0034"
/complement(5892. .6431)
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0034"
/protein_id="AAC6426.1"
/db_xref="GI:2687917"
/translation="MNKLLIPLAFECVFPSPANDSKNGAFMSAGEKLLVYETSK
ODPIYPLNLNPLGRTSGSPAGÖDLGSLITIGFPAVIGILNAGYIADITRALDSITK
KAARFÖWKGVMGLAGVYVMAVTRLTETILPTFANYSNRKLNSLNAVLAGFEPSPD
VAMGSSALGFEELSPKSY"
/complement(6499. .8379)
/gene="BB0035"
/complement(6499. .8379)
/note="similar to GP:1405462 percent identity: 31.66;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parC)"
/protein_id="AAC6419.1"
/db_xref="GI:2687910"
/translation="MDIRTVLKDNLQYSSVYIKDRAISVVDGFKPVÖRRIILSFE
MHÖGNFHVANVGMTRKYHPHÖGTISYEAIVNANKDLFTEÖGFGNLFÖDPASA
SRYTECRLLPLAFDVLKSEITITYESYDÖGNNEPLYPAKIPIVTLIOGSGIAGMA
AKIIPHNEELINAKSFLGSEYDIYDPFPGCVIÖVNEADGKYLVAKIEFTID
EKTIVIRELPREÖTESLISSTERKIRKNYIKVSIÖNPTFENAVIEISLPGVYASE
VIEKLYHTNCÖISISVNLKLLSERYPVVYTKDLIRKHAHLÖKILMELEÖSKI
LEKIEYKTLÖQIEFKKIKYKLEFITSKEENILSITSEVLNKHESFSREVLKEDENL
KIPILRKISLEÖIDNSKDIÖKILNKELESINSISISIGSYINFPÖDLLLAKSEÖHÖR
KTKISLKSÖKWEKETAÖNMVYVLAGPAGTSLÖFÖEÖIGMSYÖKYLIVPENSY
VLKNIÖEDTPIÖDKRVÖALVYDINNKSÖQISITITFNNLDNPPYÖKRRKIDKPTIDK
YVEFÖENDEVFDESINPEFESTNKDIÖVKRIÖDNEVWKSÖSSIGKRISSNMLKVK
FK"
/complement(8379. .10178)
/gene="BB0036"
/complement(8379. .10178)
/note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018
PID:467396 percent identity: 37.98; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parE)"
/protein_id="AAC6418.1"

```

Query Match 74.4%; Score 641.4; DB 1; Length 14102;  
 Best Local Similarity 87.0%; Pred. No. 1.1e-86;



JOURNAL Infect. Immun. 69 (5), 3323-3334 (2001)  
 MEDLINE 21189251  
 PUBMED 11292755  
 REFERENCE 2 (bases 1 to 777)  
 AUTHORS Noppa, L., Osterberg, Y. and Bergstrom, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden  
 FEATURES  
 source  
 1. 777  
 /organism="Borrelia afzelii"  
 /strain="ACAT"  
 /db\_xref="taxon:29518"  
 /chromosome="large linear"  
 121. 657  
 /gene="p13"  
 121. 657  
 /gene="p13"  
 /codon\_start=1  
 /transl\_table=1  
 /product="membrane protein p13"  
 /protein\_id="AAD28361.1"  
 /db\_xref="GI:4731123"  
 /translation="MNKFLIVLAFVSSFAQADSKAFNIGAEKLLAVETSK  
 DPVPLNLFGLGIGSFAOGDILGFLIGDPAVGILGAYLIDIALDNAPK  
 AAFKTMGKMLAGAVTMAVTRLEIIIFTEFANSYNRKLNINIAFGFEPSPDI  
 MNGQASLGELESEKSY"  
 BASE COUNT 228 a 84 c 148 g 316 t 1 others  
 ORIGIN  
 Query Match 88.6%; Score 764; DB 1; Length 777;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-104;  
 Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 699 AATATGGCCAAAGCTAGCGCTTGGTGGTGAACATCTTCAAAAAAGTTAATTT 758  
 Db 601 AATATGGCCAAAGCTAGCGCTTGGTGGTGAACATCTTCAAAAAAGTTAATTT 660  
 QY 759 TATTTTATTAATAAGAGATAGACAAATTTTATTTGATTCATTCATTTGA 818  
 Db 661 TATTTTATTAATAAGAGATAGACAAATTTTATTTGATTCATTCATTTGA 720  
 QY 819 AATTAGACCTTTGTTTATTTATTTATTTATTTCTCTGCTAA 862  
 Db 721 AATTAGACCTTTGTTTATTTATTTATTTATTTCTCTGCTAA 764  
 RESULT 2  
 AEO01117/c  
 LOCUS 14102 bp DNA linear BCT 15-DEC-1997  
 DEFINITION Borrelia burgdorferi (section 3 of 70) of the complete genome.  
 ACCESSION AEO01117 AEO00783  
 VERSION AEO01117.1 GI:2687907  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Borrelia burgdorferi.  
 Borrelia burgdorferi.  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia  
 burgdorferi group.  
 1 (bases 1 to 14102)  
 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,  
 Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,  
 Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D.,  
 Richardson, D., Peterson, J., Kierlavage, A.R., Quackenbush, J.,  
 Salzberg, S., Hanson, M., van-Vugt, R., Palmer, N., Adams, M.D.,  
 Goearne, J.D., Weidman, J., Uterback, T., Watthey, L., McDonald, L.,  
 Atlach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K.,  
 Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.  
 Direct Submission  
 Submitted (12-DEC-1997) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 FEATURES  
 source  
 1. 14102  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 95. 733  
 /gene="BB0027"  
 95. 733  
 /gene="BB0027"  
 /note="hypothetical protein; identified by Glimmer;  
 putative  
 1. /transl\_table=1  
 /product="B. burgdorferi predicted coding region BB0027"  
 /protein\_id="AAC66429.1"  
 /db\_xref="GI:2687920"  
 /translation="MKRYIRLIIIVALLIGVNIKIAAANDRHNSITGLISVGI  
 PIRYNDKAPRNLRYPGGIGAKIYOHILANLAIGELAYMRFIDNHSFNLDPS  
 SVGRIEYSPITTSINIFDIEGFLPIVFTNIGSLNITGDRNNITNLRTFDALPT  
 ISRSGILMNFNKKWARGATASWMMFEPNSAKMAHFALVLSIVYVNVKL"  
 741. 1790  
 /gene="BB0028"  
 741. 1790  
 /gene="BB0028"  
 /note="hypothetical protein; identified by Glimmer;"

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1923.68 Seconds

(without alignments)  
13040.962 Million cell updates/sec

Title: US-09-508-487-20

Sequence: 1 gatttcattgatccacg.....atatattatctctgcctaa 862

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl1:\*  
1: gb\_da:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_pl:\*  
25: em\_pi:\*  
26: em\_ri:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_pln:\*  
35: em\_hlg\_rtd:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vrt:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	764	88.6	777	1	AF085740	AF085740 Borrelia
2	641.4	74.4	14102	1	AE001117	AE001117 Borrelia
3	561.8	65.2	781	1	AF085739	AF085739 Borrelia
4	542.4	62.9	678	1	AF085741	AF085741 Borrelia
5	100	11.6	27323	1	AE000789	AE000789 Borrelia
6	94.4	11.0	29766	1	AE000786	AE000786 Borrelia
7	88.2	10.2	52971	1	AE001584	AE001584 Borrelia
8	71.4	8.3	43993	2	AC116965	AC116965 Dictyostel
9	67.4	7.8	14950	6	AX346159	AX346159 Sequence
10	67.2	7.8	349980	6	AX344568	AX344568 Sequence
11	67	7.8	5930	6	AX345419	AX345419 Sequence
12	67	7.8	6012	6	AX251052	AX251052 Sequence
13	67	7.8	6012	6	AX344183	AX344183 Sequence
14	67	7.8	6012	6	AX348580	AX348580 Sequence
15	66.4	7.7	129360	2	AC117079	AC117079 Dictyostel
16	65.8	7.6	9742	6	AX348912	AX348912 Sequence
17	65.4	7.6	115489	2	AC117072	AC117072 Dictyostel
18	65.2	7.6	11036	6	AX277954	AX277954 Sequence
19	65.2	7.6	11036	6	AX323649	AX323649 Sequence
20	64.8	7.5	104992	2	AC005504	AC005504 Plasmid
21	64.6	7.5	131271	2	AC015927	AC015927 Homo sap
22	64.4	7.5	6590	6	AX251927	AX251927 Sequence
23	64.4	7.5	6590	6	AX346285	AX346285 Sequence
24	64.4	7.5	6590	6	AX348730	AX348730 Sequence
25	64.4	7.5	19087	6	AX345695	AX345695 Sequence
26	63.8	7.4	17967	6	AX458472	AX458472 Sequence
27	63.8	7.4	17967	6	AX345917	AX345917 Sequence
28	63.6	7.4	349980	6	AX344565	AX344565 Sequence
29	63.4	7.4	5815	6	AX346137	AX346137 Sequence
30	63.4	7.4	16766	6	AX347059	AX347059 Sequence
31	63.4	7.4	349980	6	AX344551	AX344551 Sequence
32	63.4	7.4	349980	6	AX344552	AX344552 Sequence
33	63.2	7.3	5413	6	AX251450	AX251450 Sequence
34	63.2	7.3	7441	6	AX348445	AX348445 Sequence
35	63.2	7.3	7479	6	AX339173	AX339173 Sequence
36	63.2	7.3	13784	6	AX348449	AX348449 Sequence
37	63.2	7.3	19380	6	AX252128	AX252128 Sequence
38	63	7.3	106434	3	AC117080	AC117080 Dictyostel
39	63	7.3	163290	2	AL845514	AL845514 Danio rer
40	62.8	7.3	133501	2	AC116956	AC116956 Dictyostel
41	62.6	7.3	6716	6	AX346684	AX346684 Sequence
42	62.4	7.2	5963	6	AX281502	AX281502 Sequence
43	62	7.2	6509	6	AX251781	AX251781 Sequence
44	62	7.2	6509	6	AX344185	AX344185 Sequence
45	61.8	7.2	5884	6	AX347067	AX347067 Sequence

#### ALIGNMENTS

RESULT 1  
AF085740 LOCUS AF085740 777 bp DNA linear BCT 24-APR-2001  
DEFINITION Borrelia afzelii membrane protein p13 (p13) gene, complete cds.  
ACCESSION AF085740  
VERSION AF085740.1 GI:4731122  
KEYWORDS  
SOURCE  
ORGANISM Borrelia afzelii.  
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.  
1 (bases 1 to 777)  
REFERENCE  
AUTHORS Noppa,L., Ostberg,Y., Lavrinovitcha,M. and Bergstrom,S.  
p13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains

Pred. No. is the number of results predicted by chance to have a

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1693.82 Seconds  
(without alignments)  
13040.962 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759  
Sequence: 1 atgttaagaagaatgaat.....atctgtatgaatgggtg 759

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_bt:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sy:\*
- 12: gb\_un:\*
- 13: gb\_un:\*
- 14: gb\_vt:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vt:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rtd:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	753.2	99.2	14102	1	AE001117	AE001117 Borrelia
2	703.8	92.9	781	1	AF085739	AF085739 Borrelia
3	550.8	72.6	777	1	AF085740	AF085740 Borrelia
4	513.4	67.6	678	1	AF085741	AF085741 Borrelia
5	104.8	13.8	52971	1	AE001584	AE001584 Borrelia
6	92.6	12.2	29766	1	AE000786	AE000786 Borrelia
7	90.6	11.9	27323	1	AE000789	AE000789 Borrelia
C 8	67.8	8.9	173540	2	AC119687	AC119687 Rattus no
C 9	66	8.7	28601	1	AE000784	AE000784 Borrelia
C 10	62	8.2	236130	9	AL592044	AL592044 Human DNA
11	57	7.5	106434	3	AC117080	AC117080 Dictyoste
12	56.6	7.5	6852	6	AX251949	AX251949 Sequence
13	56.6	7.5	6852	6	AX348744	AX348744 Sequence
C 14	56.4	7.4	193379	2	AC121387	AC121387 Rattus no
C 15	56	7.4	133501	2	AC116956	AC116956 Dictyoste
16	55.6	7.3	18855	6	AX345512	AX345512 Sequence
17	55	7.2	53561	1	AE000790	AE000790 Borrelia
18	54.6	7.2	7106	6	AX346380	AX346380 Sequence
19	54.2	7.1	73334	6	AX347027	AX347027 Sequence
20	54.2	7.1	73334	6	AX356494	AX356494 Sequence
21	54	7.1	18624	6	AX346604	AX346604 Sequence
22	54	7.1	60604	2	AC023466	AC023466 Homo sapi
23	54	7.1	73377	2	AC117077	AC117077 Dictyoste
24	53.8	7.1	396	1	BORS3G	LA1151 Borrelia bu
C 25	53.8	7.1	115489	2	AC117072	AC117072 Dictyoste
C 26	53.8	7.1	124820	2	AC117073	AC117073 Dictyoste
27	53.6	7.1	33068	3	AC115682	AC115682 Dictyoste
C 28	53.4	7.0	2532	1	CJ009019	U09019 Campylobact
29	53.4	7.0	30726	2	AC117269	AC117269 Dictyoste
30	53.4	7.0	155932	1	CJ11168X6	AL139079 Campyloba
31	53.4	7.0	349980	6	AX344559	AX344559 Sequence
32	53.2	7.0	156533	2	AC117070	AC117070 Dictyoste
33	53.2	7.0	163034	2	AC099432	AC099432 Rattus no
34	53.2	7.0	349980	6	AX344566	AX344566 Sequence
C 35	53	7.0	3606	3	DD11093A	M19469 Dictyosteli
C 36	52.8	7.0	7921	3	AF153362	AF153362 Dictyoste
37	52.6	6.9	12025	6	AX346200	AX346200 Sequence
38	52.6	6.9	84805	2	AC116918	AC116918 Dictyoste
39	52.6	6.9	164399	3	PFMAL3P6	Z98551 Plasmodium
40	52.2	6.9	6012	6	AX251975	AX251975 Sequence
41	52.2	6.9	6012	6	AX344367	AX344367 Sequence
42	52.2	6.9	6012	6	AX348760	AX348760 Sequence
43	52.2	6.9	12507	6	AX345201	AX345201 Sequence
44	52	6.9	6963	6	AX345881	AX345881 Sequence
45	51.8	6.8	5586	6	AX348390	AX348390 Sequence

## ALIGNMENTS

RESULT 1  
AE001117/c  
LOCUS AE001117 14102 bp DNA linear BCT 15-DEC-1997  
DEFINITION Borrelia burgdorferi (section 3 of 70) of the complete genome.  
ACCESSION AE001117 AE000783  
VERSION AE001117.1 GI:2687907  
KEYWORDS  
SOURCE  
ORGANISM  
Borrelia burgdorferi.  
Borrelia burgdorferi  
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia  
burgdorferi group.  
1 (bases 1 to 14102)  
Fraser/C.M., Casjens/S., Huang/W.M., Sutton/G.G., Clayton/R.,  
Lathigra/R., White/O., Ketchum/K.A., Dodson/R., Hickey/E.K.,  
Gwinn/M., Dougherty/B., Tomb/J.F., Fleischmann/R.D., Richardson,D.,

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS

Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., Van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J., Venter, J.C. et al.  
Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*  
Nature 390 (6660), 580-586 (1997)

2 (bases 1 to 14102)  
Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gilm, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., Van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J.D., Weidman, J., Uitterlinden, T., Watthey, L., McDonald, L., Artlich, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K., Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.  
Direct Submission  
Submitted (12-DEC-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
Location/Qualifiers  
1. 14102  
/organism="Borrelia burgdorferi"  
/db\_xref="taxon:139"  
95. .733  
/gene="BB0027"  
95. .733  
/note="hypothetical protein; identified by Glimmer; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="B. burgdorferi predicted coding region BB0027"  
/protein\_id="AAC66429.1"  
/db\_xref="GI:2687920"  
/translation="MKRIYFIILAVLLIGVNIKKIAAANIDRHTSTGLDLSVCIPIFYNDLSKAPYTNLYPGGIGAIKIQHILNLALIGLEIKYERFIDNHSFNILNPS SVGKIFYSPIITFSINIFIDIGELFQIPVTNIGFSLNTYGDNRNNITNLRTDALPTISFGSGILMNFNKNMAFGATASWMMFEFGNSAKMAHFAVLSLVYVNVKL"  
741. .1790  
/gene="BB0028"  
741. .1790  
/note="hypothetical protein; identified by Glimmer; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="B. burgdorferi predicted coding region BB0028"  
/protein\_id="AAC66428.1"  
/db\_xref="GI:2687919"  
/translation="MKQYENYFKKRIILNLLIFLLACSSSEIFISQNLQIKIKHEYNIIGSSSPRGISLIGETLYIAAMHLPKENGKIEKIDLSNYSERINDIVNSGTYLAKNKEEELCELDGKDWTLKPKKPLKAYKELKSVGDGKVAEYIILADNNRKRIFDLOGSDKTPPOATENDKRYQISNEENLITGSLSKIMWNNNTYTRIDYQAKELPIIKTSTRGSEVAVMTGGVNNLDPKERKVSNTNNYTPITIDEGVEFSYFAEEFDALIGSNNGFAEPTKKEGIFALRAPSVEPGAVNGSOLSLGLNDIIPVSNNTYIILQGGIKLWLNKRIKKE"  
741. .2207  
/gene="BB0029"  
741. .2207  
/note="similar to SP:PI7868 GB:X03897 PID:40141 PID:1303831 GB:AL009126 percent identity: 32.58; identified by sequence similarity; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="conserved hypothetical protein"  
/protein\_id="AAC66423.1"  
/db\_xref="GI:2687914"  
/translation="MLSKIFYDADSCNFRIKPIQKPIILNKSCLTIYVSNKFFSELETTENVALEVVNDVDSFLINLADRYSLVITRIDIFPAKLILDLGKVNDEGRIFDINNINZLYFRSKINFNLIKIKRYDEDEFNKSRYEKRTIMFYSLFS"  
complement(2197. .2832)

CDS

/gene="BB0030"  
complement(2197. .2832)  
/note="similar to GB:Z36269 PID:1049113 percent identity: 30.95; identified by sequence similarity; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="signal peptidase I (lepB-1)"  
/protein\_id="AAC66422.1"  
/db\_xref="GI:2687913"  
/translation="MAPYLFEQRLRKROKRYEKFLLFLINEFFTKFVLOIPMIKSNEMPTITKNSLFEVATHLTSEFIPILKMDVIYEDPELSNFFLLIKDFEFLN KIKRRAKYKSRVIAVOGDSYVRGILVYNKDCDTEPEYINGNLVSYVLLNDPEFNDE VICEFLKNEFFLLNDNSVLNDSRIFGPIKNALVFLVKVVDKIKY"  
complement(2834. .3814)  
/gene="BB0031"  
complement(2834. .3814)  
/note="similar to PID:1652260 percent identity: 35.98; identified by sequence similarity; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="signal peptidase I (lepB-2)"  
/protein\_id="AAC66421.1"  
/db\_xref="GI:2687912"  
/translation="MYIRLDKFPASFLVHVEKYLYYRRKKRYECKLRKRRGFLNLFDFVAALFVLVINOYEVQAKIPSGSMENLTQIDDFLVDFVDFSYGPELPELFIN GFKAPESDIIIFENPEYKSGVFPDIFORILYMLTSLFDLDRDYGNNVFLVKGAPADKIVRNNNGKAVIKREGSENFLEDSYDLVDKNEKIKKIYANEDYGIYGDPA MPVALSOLNINLSSTPDFSFDFRMDIDREFEELKYSAPFMDYVYKEMIMBDYGIYVPGDYLIPIDGNRDNHSDGRFPGYINKKVKGLRIIYLPFSVGF"  
3885. .3357  
/gene="BB0032"  
3885. .3357  
/note="hypothetical protein; identified by Glimmer; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="B. burgdorferi predicted coding region BB0032"  
/protein\_id="AAC66427.1"  
/db\_xref="GI:2687918"  
/translation="MKREVDSNVELFEVYKPIIGTEPVYVFLITILLLSLSTLIIIPKFNPGAVLAKINSINENTYIYLEKTIQRTPLKRYINATEGVIRAKRMKPYTE QRIKIHKKFPGNYSLOINLELVDEKTIKQROKELISWIKININETKLIIPVLSLS SELKEHPKYIKKEFKDSIPYLNSTEMKDYKALYSIDQNSQDELTMSLNKLNFDLENRAIFWFLNLDKDLITKNAAPVKTLLADNEDOLSKNEKINIKLPGEFK INSKTIKIONYELNSDKSNISLSKSYNKEPLIOQNTYKKEODFLKENKMAIN KNEKILKEQLVDENYLKNPQICLNPAITGTSYFSAFVANNYSKILPGEFKAPLIS QEWELVKEPKNKPNLNTNISKVGVGNLMQNSPFEIATFKNEKNFYSENSNFSILI TEIRTSQOQNNNLNASTKASFLKMWSSPNIQRLIVSKP"  
5369. .5821  
/gene="BB0033"  
5369. .5821  
/note="similar to GP:2073503 percent identity: 47.68; identified by sequence similarity; putative"  
/codon\_start=-1  
/transl\_table=-1  
/product="small protein (smpB)"  
/protein\_id="AAC66420.1"  
/db\_xref="GI:2687911"  
/translation="MTNNNLLLENKRAKPNYFTEBKISGIVLKGTEVKSIAKKLSF NNSPASIKKEEIMLENHVSKEGNFENDELRLLKRELQALKKFEKEGYT LIPISFLKSIIVGVGICGKKLYKRELKKSIAKKOLSRIRIKY"  
complement(5892. .6431)  
/gene="BB0034"  
complement(5892. .6431)  
/note="hypothetical protein; identified by Glimmer; putative"  
/codon\_start=-1  
/transl\_table=-1

CDS

gene

CDS

gene

Oy	421	GCTTACTTGGGGGGCTTATTTTGGATATCAACAGCGTGATGGTATTACTAAAAAGC	480
Dd	6180	GCTTACTTGGGGGGCTTATTTGGAATATCAACGCGTTGATGATTACTAAAAAGC	6121
Oy	481	TGCTTTCAATGACTTGGGGTAAAGGAGTATATGTAGCAGGTGNGGTTACTATGCGTGT	540
Dd	6120	TGCTTTCAATGACTTGGGGTAAAGGAGTATATGTAGCAGGTGNGGTTACTATGCGTGT	6061
Oy	541	GACAAGATTAAACAGAAATATTCTTCCATTACATTGCTAATTAATATAGACGCT	600
Dd	6060	GACAAGATTAAACAGAAATATTCTTCCATTACATTGCTAATTAATATAGAGAGCT	6001
Oy	601	AAAAAATAGCCTTAATGTAGCTTTTGGAGAGATTACCTAGTTTGGATGTCGAATGG	660
Dd	6000	AAAAAATAGCCTTAATGTAGCTTTTGGAGAGATTACCTAGTTTGGATGTCGAATGG	5941
Oy	661	CCAAATCAGTGTCTTGGGTTGAAGCTGCTTTCAAAAAAAGCTATTAATTTATTTAT	720
Dd	5940	CCAAATCAGTGTCTTGGGTTGAAGCTGCTTTCAAAAAAAGCTATTAATTTATTTAT	5881
Oy	721	ACAAAAATGGTGATTGCAATTTCTGTAATGAAATGGCT	758
Dd	5880	ACAAAAATGGTGATTGCAATTTCTGTAATGCAATTTGCT	5843
RESULT 2			
LOCUS	AF085739	781 bp	DNA
DEFINITION	Borrelia burgdorferi membrane protein p13 (p13) gene, complete cds.		
ACCESSION	AF085739		
VERSION	AF085739.1	GI:4731120	
KEYWORDS			
SOURCE			
ORGANISM	Borrelia burgdorferi.		
	Borrelia burgdorferi		
	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia		
	burgdorferi group.		
REFERENCE	1 (bases 1 to 781)		
AUTHORS	Noppa, L., Ostberg, Y., Lavrinovich, M. and Bergstrom, S.		
TITLE	P13, an integral membrane protein of Borrelia burgdorferi, is		
JOURNAL	C-terminally processed and contains surface-exposed domains		
MEDLINE	Infect. Immun. 69 (5), 3323-3334 (2001)		
PUBMED	11292751		
REFERENCE	2 (bases 1 to 781)		
AUTHORS	Noppa, L., Ostberg, Y. and Bergstrom, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea		
	University, Umea S-901 87, Sweden		
FEATURES	Location/Qualifiers		
source	1..781		
	/organism="Borrelia burgdorferi"		
	/strain="B31"		
	/db_xref="taxon:139"		
	/chromosome="large linear"		
gene	122..661		
	/gene="p13"		
CDS	122..661		
	/gene="p13"		
	/note="similar to Borrelia burgdorferi predicted coding		
	region BB0034 encoded by GenBank Accession Number		
	AE001117"		
	/codon_start=1		
	/translation="MKNLIFVLTATPCVFSEFAQDANKGAFMSAGEKLTLYETSK		
	ODPVPFLNIFLGFGLGSPAGQGIILGSGILIGPDVIGILLAGAYLDIKADGTRK		
	KAFQRTWKGKVMLAGYVTAATVTLTLLIPFTTANSYNKKLNSLVNLGGEPSFD		
	VAMQSSALGEUSFKRSY"		
BASE COUNT	223 a	98 c	153 g
ORIGIN			307 t

Query Match 92.9%; Score 705.2; DB 1; Length 781;  
 Best Local Similarity 99.6%; Pred. No. 5.1e-109;  
 Matches 707; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 TCTTTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAACCTTAA 108  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 TCTATTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAACCTTAA 60

QY 109 AAACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 168  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 AAACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 120

QY 169 TATGAATPAACCTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCA 228  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 TATGAATPAACCTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCA 180

QY 229 AGCTAATGATCTAAAAATGCTGGCTTGGGATGAGTGTGAGAAAAACCTTTGGTTTA 288  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 AGCTAATGATCTAAAAATGCTGGCTTGGGATGAGTGTGAGAAAAACCTTTGGTTTA 240

QY 289 TGAAGTACGACAGAGATCTATGACATTTTATGAACTTTTATAGGGTTTG 348  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 TGAAGTACGACAGAGATCTATGACATTTTATGAACTTTTATAGGGTTTG 300

QY 349 AATAGCTCCTTTGCTCAAGAGATATTCTTGAGAGTTCTCTATCTTGAGATTGATGC 408  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 AATAGCTCCTTTGCTCAAGAGATATTCTTGAGAGTTCTCTATCTTGAGATTGATGC 360

QY 409 GGTGGTATAGGGCTTATACCTTGGGGGGCTTATTTGGATTCAAAAGGGCTGATNGTAT 468  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 GGTGGTATAGGGCTTATACCTTGGGGGGCTTATTTGGATTCAAAAGGGCTGATNGTAT 420

QY 469 TACTAAAAAGCTCTTTTCAATGACCTTGGGTAAGGAGCTTATGTTAGCAGGTGTGCT 528  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 TACTAAAAAGCTCTTTTCAATGACCTTGGGTAAGGAGCTTATGTTAGCAGGTGTGCT 480

QY 529 TACTATGGCTGTGACAGATTAACAGAAATTAATTTCTTCATTACATTTGCTTAATAGTTA 588  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 481 TACTATGGCTGTGACAGATTAACAGAAATTAATTTCTTCATTACATTTGCTTAATAGTTA 540

QY 589 TAAATAGAGAGTAAAAAATACCTTAATGATAGCTTTAGAGAGATTGAACTAGTTTGA 648  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 TAAATAGAGAGTAAAAAATACCTTAATGATAGCTTTAGAGAGATTGAACTAGTTTGA 600

QY 649 TGTTCAGATGGGCAATCATCAGTGTCTGTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 708  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 TGTTCAGATGGGCAATCATCAGTGTCTGTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 660

QY 709 ATTTTATTTATTAACAAAAATGGGTGATTGCAATTCCTGATTGAATGGGT 758  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 ATTTTATTTATTAACAAAAATGGGTGATTGCAATTCCTGATTGAATGGGT 710

RESULT 3  
 AF085740 777 bp DNA linear BCT 24-Apr-2001  
 LOCUS AF085740  
 DEFINITION Borrelia afzelii membrane protein P13 (P13) gene, complete cds.  
 ACCESSION AF085740.1 GI:4731122  
 VERSION AF085740.1 GI:4731122  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Borrelia afzelii.  
 Borrelia afzelii  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia  
 burgdorferi group.  
 1 (bases 1 to 777)  
 AUTHORS Noppa,L., Ostberg,Y., Lavrinovicha,M. and Bergstrom,S.  
 TITLE P13, an integral membrane protein of Borrelia burgdorferi, is  
 C-terminally processed and contains surface-exposed domains  
 JOURNAL Infect. Immun. 69 (5), 3323-3334 (2001)  
 MEDLINE 21189251  
 PUBMED 11292755  
 REFERENCE 2 (bases 1 to 777)

AUTHORS Noppa,L., Ostberg,Y. and Bergstrom,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-Aug-1998) Microbiology, Medical Microbiology, Umea  
 University, Umea S-901 87, Sweden  
 FEATURES  
 source  
 1..777  
 /organism="Borrelia afzelii"  
 /strain="ACAT"  
 /db\_xref="taxon:29518"  
 /chromosome="large linear"  
 /gene="p13"  
 121..657  
 121..657  
 /gene="p13"  
 /codon\_start=1  
 /transl\_table=11  
 /product="membrane protein P13"  
 /protein\_id="AAD28361.1"  
 /db\_xref="GI:4731123"  
 /translation="MNKFLIVLAFCVSSRFQADDSKSFNLAGKELIAYETSK  
 DPLVPLNLFLFGIGSFAQGDILGFLILGDAVIGLILTGATLIDKADKAPK  
 AAFKWTWGMKMLAGAVTMAVVRLEIILIPFPANSYNRKLNSLNIAAGGEPSPDI  
 NMQASALGELISFKRSY"  
 BASE COUNT 228 a 84 c 148 g 316 t 1 others  
 ORIGIN

Query Match 72.6%; Score 550.8; DB 1; Length 777;  
 Best Local Similarity 87.9%; Pred. No. 4e-83;  
 Matches 624; Conservative 0; Mismatches 82; Indels 4; Gaps 2;

QY 50 CTATTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAACCTTAA 109  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 CTATTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAATTAACCTTAA 60

QY 110 AACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 169  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 AACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 120

QY 170 ATGAATAACTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCAA 229  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 ATGAATAACTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCAA 180

QY 230 GCTAATGATTTAAAAATGTCGCTTGGGATGAGTGTGAGAAAAACCTTTGCTTAT 289  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 GCTAATGATTTAAAAATGTCGCTTGGGATGAGTGTGAGAAAAACCTTTGCTTAT 237

QY 290 GAACTGACAGAGATCTATGATGACATTTTATGAACTTTTATAGGCTTTGA 349  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 238 GAACTGACAGAGATCTATGATGACATTTTATGAACTTTTATAGGCTTTGA 297

QY 350 ATAGGCTCCTTTGCTCAAGAGATATCTTGGAGGTTCTCTATCTTGATTTGATGC 409  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 298 ATAGGCTCCTTTGCTCAAGAGATATCTTGGAGGTTCTCTATCTTGATTTGATGC 357

QY 410 GTTGGTATAGGGCTTATATCTGCGGGGCTTATTTGGATATCAAGCGCTGATGTAT 469  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 358 GTTGGTATAGGGCTTATATCTGCGGGGCTTATTTGGATATCAAGCGCTGATGTAT 417

QY 470 ACTAAAAAAGCTGCTTTTCAATGAGACTTGGGTAAGGAGTATGTTAGCAGGTGTG 529  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 418 GCTCAAAAAGCGCTTTTAACTGAGCTTGGGTAAGGAGTATGTTAGCAGGTGTG 477

QY 530 ACTATGCTGTGACAGAGATTAACAGAAATTAATCTTCATTACATTTGCTTAATAGTTAT 589  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 478 ACTATGCTGTGACAGAGATTAACAGAAATTAATCTTCATTACATTTGCTTAATAGTTAT 537

QY 590 AATAGGAGCTAAAAAATAGCCTTAATGTACCTTAGAGGATTTGAACTAGTTTGA 649  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 538 AATAGGAGCTAAAAAATAGCCTTAATGTACCTTAGAGGATTTGAACTAGTTTGA 597

QY 650 GTTGCATGGGCAATCCAGTGTCTTGGGTTTGAACGTCTTTCAAAAAAGCTATTTAA 709  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 598 ATTATATAGGCAAGCTAGCGCTTGGGTTTGAACATATCTTTCAAAAAAAGTTATTTAA 657

```

OY 710 TTTTA-TTATATCAAAATGGGTGATTCATTCGTATGGAATGGG 758
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 658 TTTTATTTATTTATTAATGAGTGAATTTGTATGTGATTGCT 707

RESULT 4
AF085741 678 bp DNA linear BCT 24-APR-2001
LOCUS AF085741
DEFINITION Borrelia garinii membrane protein p13 (p13) gene, complete cds.
ACCESSION AF085741
VERSION AF085741.1 GI:4731124
KEYWORDS
SOURCE
  Borrelia garinii.
  Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
  burgdorferi group.
REFERENCE
  1 (bases 1 to 678)
  Noppa, L., Ostberg, Y., Lavrinovitch, M. and Bergstrom, S.
  p13, an integral membrane protein of Borrelia burgdorferi, is
  C-terminally processed and contains surface-exposed domains
  Infect. Immun. 69 (3), 3323-3334 (2001)
JOURNAL MEDLINE 21189251
PUBMED 11292755
REFERENCE 2 (bases 1 to 678)
AUTHORS Noppa, L., Ostberg, Y. and Bergstrom, S.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea
University, Umea S-901 87, Sweden
FEATURES
  source
    1..678
    /organism="Borrelia garinii"
    /strain="Ip90"
    /db_xref="taxon:29519"
    /chromosome="large linear"
    /gene="p13"
    /gene="p13"
    /gene="p13"
    /codon_start=1
    /transl_table=11
    /product="membrane protein p13"
    /protein_id="AAD28362.1"
    /db_xref="GI:4731125"
    /translation="MNKRLIFLIVFCAPFSRQDDSKSTFNLGGEKFLVETNKKD
    SVLPFLNLFJGFGISRAQGDILGSDILGDAIGLILGALDILKDEPNNAKKA
    DFKMTWKGMLAGVYMAVTRLEIVLPFTFANYNRKLNLSINIALAGPEPFDIN
    MGOASALGFGLSFKFSY"
  gene
    121..654
    /gene="p13"
  CDS
    121..654
    /codon_start=1
    /transl_table=11
    /product="membrane protein p13"
    /protein_id="AAD28362.1"
    /db_xref="GI:4731125"

BASE COUNT 202 a 77 c 140 g 259 t
      RIGIN

Query Match 67.6% Score 513.4, DB 1; Length 678;
Best Local Similarity 86.4%; Pred. No. 7.9e-77;
Matches 592; Conservative 0; Mismatches 86; Indels 7; Gaps 2;

```

```

Db 234 TGAACATAAAGAAAGATTCCTGTACCAATTTTATGAACTTTTATGAGGTCGG 293
OY 349 AATAGCCTCCTTTCCTCAAGAGATATTCCTGAGGCTCTTATTCCTGGATTATGCG 408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 294 GATAGGCTCTTTCCTCAAGAGATATTCCTGAGGCTCTTATTCCTGGATTATGCG 353
OY 409 GGTGGTATAGGCTATATCTGGCGGCTTATTTGAGATACCAAGCCGTGATGCTAT 468
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 GGTGGTATAGGCTATATCTGGCGGCTTATTTGAGATACCAAGCCGTGATGCTAT 413
OY 469 TACTAAAAAGCTCTTTCATGAGACTTGGGTAAGGAGATTTGATGACAGTGTGGT 528
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 TGCTAAAAAGCTGATTTTAAGTGACCTTGGGTAAGGAGATGCTTGGCAGCTGTGG 473
OY 529 TACTATGGCTGTGACAAATTAACAGAAATATTCCTCAATTTACATTTGCTATATG 588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 TACTATGGCTGTGACAAATTAACAGAAATATTCCTCAATTTACATTTGCTATATTA 533
OY 589 TATATAGAGCTAAAAATAGCCTTAATGATAGCTTAGAGGATTTGAACCTAGCTTTGA 648
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 TATACAGAAAGCTGAAAAATATGCTTAATATATAGCTTGGAGAGATTTGACGCTATGTA 593
OY 649 TGTTCGAATGGGCCAATTCAGTCTCTTGGGTTTGAACCTGCTTTCAAAAAAGCTATTA 708
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 TATTAACATGGGCGCAAGCTAGTCTCTTGGTGTGACGTCTTTCAAAAAAGCTATTA 653
OY 709 ATTTTATTTATTAACAAAAATGGGTG 733
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 ATTTTATTTATCTAGAAAAATGGGTG 678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
AE001584 52971 bp DNA linear BCT 21-JUN-2002
LOCUS AE001584
DEFINITION Borrelia burgdorferi plasmid lp56, complete plasmid sequence.
ACCESSION AE001584
VERSION AE001584.1 GI:6382393
KEYWORDS
SOURCE
  Lyme disease spirochete.
  Borrelia burgdorferi
  Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
  burgdorferi group.
REFERENCE
  1 (bases 1 to 52971)
  Casjens, S., Palmer, N., van Vugt, R., Huang, W.M., Stevenson, B.,
  Rosa, P., Lathigra, R., Sutton, G., Peterson, J., Dodson, R.J., Haft, D.,
  Hickey, E., Gwinn, M., White, O. and Fraser, C.M.
  A bacterial genome in flux: the twelve linear and nine circular
  extrachromosomal DNAs in an infectious isolate of the Lyme disease
  spirochete Borrelia burgdorferi
  Mol. Microbiol. 35 (3), 490-516 (2000)
JOURNAL MEDLINE 20138354
PUBMED 10672174
REFERENCE 2 (bases 1 to 52971)
AUTHORS Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Lathigra, R.,
White, O., Dodson, R., Hickey, E.K., Gwinn, M., Peterson, J.,
van Vugt, R., Palmer, N., Haft, D., Rosa, P. and Stevenson, B.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
  source
    1..52971
    /organism="Borrelia burgdorferi"
    /db_xref="taxon:139"
    /plasmid="lp56"
    /gene="BB001"
    /gene="BB001"
    /note="similar to GB:U03641 PID:456212 percent identity:
    67.53; identified by sequence similarity; putative"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"

```

gene  
CDS  
/protein\_id="AA07704.1"  
/db\_xref="GI:6382394"  
/translation="MTRKYLIAVARRKDAEICISGVRRKQNTENKLIIVFLLCLE  
GYLGVHRYVGKIGTGLYLFEGFLYGVLDLIRITTNKFKCN"  
710..802  
/gene="BB002"  
710..802  
/gene="BB002"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA07738.1"  
/db\_xref="GI:6382428"  
/translation="MRGKIGICLVLEKSIIVGLTMFKVIR"  
856..1407  
/gene="BB003"  
856..1407  
/gene="BB003"  
/note="similar to GB:L31615 GB:L31418 PID:520783  
PID:551742 percent identity: 97.81; identified by sequence  
similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="outer membrane protein, putative"  
/protein\_id="AA07705.1"  
/db\_xref="GI:6382395"  
/translation="MRLVGVFIIALALGCVLPDNOQAVOTFPFSSISIMGSDE  
IVREGFSSKIVASEHRLVETIKTLISKDPYRDPVPRDYNEEYFKFLDGS  
EÖSKDLIKRGYRKNQNKFKSEVYWLISCISELSPDIKISGEESPEYDRMPAP  
TAYQYLVKVRKEIRNTNGKIL"  
1404..2268  
/gene="BB004"  
/note="This region contains an authentic frame shift and  
is not the result of a sequencing artifact: similar to  
GP:520778 percent identity: 93.50; identified by sequence  
similarity: putative; protein p23, authentic frameshift"  
2744..3496  
/gene="BB005"  
2744..3496  
/gene="BB005"  
/note="similar to GP:1553115 percent identity: 53.30;  
identified by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="antigen, p35, putative"  
/protein\_id="AA07706.1"  
/db\_xref="GI:6382396"  
/translation="WKYICVCFLLINACNSDSFSTNOEDIKYPSDEKSKSMSEAS  
KEEDPNKIRKNTLNDLINLEIANEKREYERKQDEPSDYGILAFQELDSVGKI  
SEDPQSKFRKNTYSPLSAIDVNRKLDLSIINSGIOGLFENRFGIIPDSIN  
HVSXKDILGGLTDLKLSNPEKLSLKEPFSKMLNOLLDDYKNDKHIRETNK  
LKSHTALFPOLDKKEDEAYEPKMQIFSISNL"  
3623..4108  
/gene="BB006"  
3623..4108  
/gene="BB006"  
/note="similar to GP:2239281 percent identity: 79.75;  
identified by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA07707.1"  
/db\_xref="GI:6382397"  
/translation="MQLFAOEKLEKGVGDIATVMKYESKRAITLAPLINTFSLIG  
SFVQGVITGGALLGSVGLIMMAGYMGWDIATVYKESRAITVITKNSIGTGLTA  
ASTTGTIIPKFAKRYNADLKRLGIALAGLEPNFDIIGINSDSNRLKNPIETQFTW  
G"  
complement(4249..4986)  
/gene="BB007"  
complement(4249..4986)  
/gene="BB007"  
complement(4249..4986)  
/gene="BB007"

gene  
CDS  
/note="similar to PID:1143400 percent identity: 62.50;  
identified by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA07708.1"  
/db\_xref="GI:6382398"  
/translation="MCKKYNKKELVANSRFTSMNKLINIDNMOKDLINYNLKEK  
LYNLKDDIDNKEKMLIYEIKKELFYDNSFEGPTKHFVIAKTAIYLKAYE  
KYLEGASIDINKETGFRGLQRLKENLNVKKNLSEVNNKMSIRFPMKNEFTYE  
CKEDIRACIYLIERLEFFTKREVLDLINEEKYKRRKKNKYLIFIIYGAIVILGMM  
MSFVLPOLSQSKRPFMGLEFSLLL"  
complement(5069..5830)  
/gene="BB008"  
complement(5069..5830)  
/gene="BB008"  
/note="similar to GB:D26185 SP:P37522 PID:467381  
PID:580906 GB:AL009126 percent identity: 54.17; identified  
by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="plasmid partition protein, putative"  
/protein\_id="AA07709.1"  
/db\_xref="GI:6382399"  
/translation="MDRKESNITITASPRGVGKTTLTIFSYILKDSKVLIDLD  
PÖNSLSYFSKYTFNDKNSYSLKLDVYFGCCIKINDFISILRSHPLENSEN  
LNYKLDLEMLINRNTNVEDYLLDTPYFLLKNSLNVSDYIITIPÖVERSE  
LSLIÖMTINDIDKDFNRKSPNISIIENÖFKNNTPEKDEELLKYEFYKGIHYY  
NSVAVLINELLESEVETEEETKNTLEINLH"  
complement(5803..6339)  
/gene="BB009"  
complement(5803..6339)  
/gene="BB009"  
complement(5803..6339)  
/note="similar to GB:U03641 PID:458217 percent identity:  
58.62; identified by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA07710.1"  
/db\_xref="GI:6382400"  
/translation="MNNTERFNLKKTKKRIFFYKIEKENKKTYSIIFKHLIK  
EITRKGLRITFOKENDGWNPFNLEPKDEKFLIKYGMNLKPEFLIGNNIKY  
VTRKAYIEVFEKKSITKCVÖSLMTLLREKKÖTYKRYKTLEINIMKEVTEFYNK  
KIXEBGIINKWIKNÖTL"  
complement(6336..6674)  
/gene="BB010"  
complement(6336..6674)  
/gene="BB010"  
complement(6336..6674)  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA07739.1"  
/db\_xref="GI:6382429"  
/translation="MYNSESRAVALIANGPPALITPLFINKDFKHLKYÖY  
TKDKEFLYKLLKKNKIHFMÖNÄHETDFFRYIGEEFÖTIIKKIKINÖNTFLFD  
LÄNYKRRKI"  
6800..7411  
/gene="BB012"  
6800..7411  
/gene="BB012"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA07740.1"  
/db\_xref="GI:6382430"  
/translation="MVRPMAIKGVQÄADTTÖLLKDYÖDRRSKLEKFMKPÖYDAGLL  
SNSVEFRDKNIÖEFASGTRTSKDLNHPFSGYPKRKIKRYIOEKADÖIHIEPH  
VEAGGEDLVGICIDIDEFSKTATIVITNNFEGYLVAKSDTLKVKDKLVFNKDGAL  
KYGAPPNKATINAIASDAKÖISNDVYLLIKAVFNGKAVSKN"





```

putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG04"
/protein_id="AAC66074.1"
/db_xref="GI:2690031"
/translation="MKQEAISIFDLIAVHIKTLIDINLNIYNGVV"
/complement(2891..4056)
/gene="BBG05"
/notes="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
GP:228714 percent identity: 99.12; identified by sequence
similarity; putative; transposase-like protein, authentic
frameshift"
4208..5368
/gene="BBG06"
4208..5368
/notes="BBG06"
/notes="similar to GB:003641 PID:458216 percent identity:
66.76; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66054.1"
/db_xref="GI:2690011"
/translation="MKRVFTFLKLCITINIPRSSTMTNSKPPCHNKLOQLIV
LSTLAVNSKYNKYTKQNLICENMLKRGQPTTLTQNTLYLEKFKYNTY
YKHLINNGTEIYKLYNKSKKCYLKHQREKDYRFKADVDNYLNDKNGKSYD
LVLECLNNNNIKERKIIOEKYQVAKFNKQFSLKELPLINDIKDELKIL
KIIRIEINLTNNNTYLYNYPEKONKLLKLSNTQOOLEKNGYPEOLEISFOKI
YENKTPHPIENHKYDLSHKRLEKSTERKENLQDCENNRINIFITLQOLK
KEYIDVLPKLIKLYNSKNKLEYNKVNMYELLEIIKKEKNOLKEVY"
5378..5950
/gene="BBG07"
5378..5950
/notes="BBG07"
/notes="similar to PID:114398 PID:1174342 PID:1174350
PID:2182755 percent identity: 57.84; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66055.1"
/db_xref="GI:2690012"
/translation="MQMMAKSIOQLKPIVSCNKKDLFIKIEKNDKTIYTKIMDI
YKFLINKKKNKRRISLGLFQNSKVEEFNLADPGLGIIYGYKRIKIKVYO
VGGIEKSYLSKTYMEVRFKSGIIFCYFSLPILKENVNTYNNALFMSMTLLER
VGYDFYKXKPKQKPFIKWIKESWLNKHI"
5911..6678
/gene="BBG08"
5911..6678
/notes="BBG08"
/notes="similar to PID:1174351 percent identity: 39.26;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="stage 0 sporulation protein J (spoJ)"
/protein_id="AAC66056.1"
/db_xref="GI:2690013"
/translation="MDRKLAKRPHIIAISIKGVGKSTSTMESTILSTNKVLV
DLDQNAVTSYFITQDHPRMELINITYSILKHKFKDVVISISKNLDFISYLEL
AKTSKGNQKREMLRNAVNYNLEDYVITDPPSSSELDNALVADAVIIPVLE
RMAVENPLDINQIKELNNFMGEAKIHFASKVIGISGRTSTEISMLKEKYNLF
IGEVHSEALKIIDVAIGPENENYKKEVLRILEKI"
6737..7288
/gene="BBG09"
6737..7288
/notes="BBG09"
/notes="similar to PID:1143400 percent identity: 43.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

```

```

/protein_id="AAC66057.1"
/db_xref="GI:2690014"
/translation="MKNRKVIINDIRIVNATYINTEERKKEELKNEKNRIED
IRKNTMKILLETIRNKLIIDIOYKRFEPFIDPKIARQAKYIKIATIEGKLE
EIDITENGIDRTLEMLNKKRKINSKANLITPLRILETOACDFYKMPNFANTILED
FYOKNEQOLIKLEEKYKOKYS"
/complement(7483..10779)
/gene="BBG10"
/complement(7483..10779)
/notes="BBG10"
/notes="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG10"
/protein_id="AAC66075.1"
/db_xref="GI:2690032"
/translation="MKIDETIPLMSISNNOKIDAETLTKIADOKESSLDDLSK
LEKSAKSGNMLEKALNKSQNAIKNFSLANSVNSVSKSGMSKSTICKYLSVKGIG
SVKMLANKTEAEFQOMIAAPILIMAVKAIQAGTISGIFDGMADLDFNEVSF
SDMLGNEDLKSIAESMRARFDETLFRDAIIVATKMLSYGATASEBERIRMGFA
AGSSSEGLEKLAEYSSVESNOVNLEDIALDAGVDITDILAEERGLGALYKRA
SDKIGSDALNKALSKATSESGCKEYGTAKERTIAAOOCTAKMSKRLPLDGLALE
PMKIGPKYKQFLMVAIMVPLTKYTTITITLTTKIGELVILAGKYVQKMKFKALT
TPEVKIEGIRKVIYGLKLELATYVSSGFLDFRFGAEPILIKWQKLDIMTSKAYTLK
NLVIFWRKAERKEDTSGSEPERKGPDPNKTINIKMAEDYOKLDEIPRORDIYN
KTGARQOALRNLEKTIENKQKFKDEYSKIFPOLDENKILIVGEKSYNERNNSY
DEVNEYONLKEKRSERETIKTLPHRDQVSALOKLNDEINERKMAFEVYKGFETL
NESRQVYNLKEOVNEYETKLDPSVEVLEOKLQKIDPLEMETMLPKREBASAEK
KMAKIDAMKRYEYDEKRSQPKLINEINRTIQOYAKADDTKSLSDMSIDGLNFK
NAFKEDTAGFVLNDGTESIGEERHNILNKRDVNWGSKMTOTMESKTKITPA
GAVFGPEAVFLEINGLDFVWGLKNGOKARKALEKRDDBLELEKRESELEK
LEDFDEEIKMRKEKSELDEYTKLEIFLOKQASGOISGEERFOKLADVOYETKTK
KDIEFTQITKEAKKLEVERKNNKLSLEPERIAKQAEVDKNAEDMMNGLKEDRLK
TOKIDELIRIKRVKSGSIEPEKLAHGARFVSNKPYMPNSGVSSPFGQPELVR
ITPAIDENIRKLEAKITIAEITTKLOQTQSSIVNNFYTFNGDVLDAEKLVRMLSK
EHLMGFRMAE"
/complement(10776..11015)
/gene="BBG11"
/complement(10776..11015)
/notes="BBG11"
/notes="hypothetical protein; identified by Glimmer;
putative"

```

Query Match 12.2%; Score 92.6; DB 1; Length 29766;  
Best Local Similarity 52.1%; Pred. No. 1.1e-06;  
Matches 262; Conservative 0; Mismatches 229; Indels 12; Gaps 2;

QY	165	TTTTATGATAAAGCTTTTAAATTTTGTGGACACCTTTGTGTTTCTAGCTTTG	224
DB	2099	TTAATATGAAAAAATTTTCAATTAATTAATTTAGTTTAAACAATCCTTTG	2158
QY	225	CTCAAGCTAATGATCTTAATAATGCGTTGGAGTAGTGGAGAAACCTTTGG	284
DB	2159	CACAAGATAAAGCTTGAANA-----GCTGGAGTGGAGATATCGCGCTAATG	2211
QY	285	TTTATGAACATAGACAGCAAGATCCTATTGACCACTTTTATGAACCTTTTAA	344
DB	2212	AATATGAAGAAGGAAAAAACAACCACTAGACACATCTTTGAAATATCTTTTGT	2271
QY	345	TTGGAATAGGCTCTTTCTCTCAAGAGATATCTTGGAGTTCTTATTTCTTGG	404
DB	2272	TAGGATAGGATCTTTTCTTCAAGAGATATATATGAGTGAGTCACTATGCTCT	2331
QY	405	ATGCGGTGATATAGGCTTATCTTGGGGGCTTATTGGATATCAAAAGCGTTATG	464
DB	2332	AAATGCTTGGAA-----GGAATACTTAATATGCGTGGACATGATCGTAGAG	2386
QY	465	GATATTAATAAAGCTCTTTCAATGAGCTGGAGTAAGGAGATATTTAGACAGT	524
DB	2387	GAGCGTTACAGAAAGCAGACACAGTAATATACCGAGGAGTATATACAGAAAT	2446
QY	525	TGTTACTATGCGTGTGACAAGATTAAACGAATTAATCTTCATTTTACATTTCT	584

[illegible]

```

gene      /gene="BB108"  
CDS       3576..3755  
          /note="hypothetical protein; identified by Glimmer;  
          putative"  
          /transl_table=1  
          /codon_start=1  
          /product="B. burgdorferi predicted coding region BB108"  
          /protein_id="AAC66196.1"  
          /db_xref="GI:2690093"  
          /translation="MISFDYNPASNFIKITTPCKLAKLLIPLIOKAQFNLES  
          FYKNIKYSRSLNLIGK"  
          /gene="BB109"  
          3745..3882  
          /note="hypothetical protein; identified by Glimmer;  
          putative"  
          /codon_start=1  
          /transl_table=1  
          /product="B. burgdorferi predicted coding region BB109"  
          /protein_id="AAC66197.1"  
          /db_xref="GI:2690094"  
          /translation="MASEAKIDAESKEVNSOKSKNSKNIPEKDTGLIVAKIRKS  
          L"  
          3911..4315  
          /gene="BB110"  
          3911..4315  
          /gene="BB110"  
          /note="hypothetical protein; identified by Glimmer;  
          putative"  
          /codon_start=1  
          /transl_table=1  
          /product="B. burgdorferi predicted coding region BB110"  
          /protein_id="AAC66198.1"  
          /db_xref="GI:2690095"  
          /translation="MGRAMTPHQQLKYVLENGKLMLDLDELKTAAIDGDNNNISKY  
          NDLVKAVDPENNNOSLIKVLYLKSSNEEDIKAKCICITLMHNVETFEVSODELDKRNK  
          KNEHSILTTLEALGRIKNSAIAIHVCFNID"  
          complement(4623..4721)  
          /gene="BB111"  
          complement(4623..4721)  
          /gene="BB111"  
          /note="hypothetical protein; identified by Glimmer;  
          putative"  
          /codon_start=1  
          /transl_table=1  
          /product="B. burgdorferi predicted coding region BB111"  
          /protein_id="AAC66199.1"  
          /db_xref="GI:2690096"  
          /translation="MSKLGSTVNYPVKTIKILLYNVKGSYLK"  
          5128..5346  
          /gene="BB112"  
          5128..5346  
          /gene="BB112"  
          /note="hypothetical protein; identified by Glimmer;  
          putative"  
          /codon_start=1  
          /transl_table=1  
          /product="B. burgdorferi predicted coding region BB112"  
          /protein_id="AAC66200.1"  
          /db_xref="GI:2690097"  
          /translation="MIKEPKNHISKIHIIKEVEDVTIKWDREYSLFKLHGKNTLED  
          WLEYTGKEKNQIKIEFANKFIKKRRPKI"  
          5609..5707  
          /gene="BB113"  
          5609..5707  
          /gene="BB113"  
Query Match    11.9%; Score 90.6; DB 1; Length 27323;  
Best Local Similarity 50.4%; Pred. No. 2.5e-06;  
Matches 325; Conservative 0; Mismatches 289; Indels 31; Gaps 3;
```

OY	24	TATTTTATTGTCGAATCAAGAAGCTCATNTGGAGAGGAAATTTCCAACCAATAATTTCGA	83
Dd	20256	TATATTTTAAAACCAAAAPAAACCTAATTTTAAACAACCTCATTAAACCCAAATTTTCATA	20197
OY	84	AAAACTTAATTAATAA-ACCTTA AAAAACCTTTTTTAATTCATTAATATGTACCAT	142
Dd	20196	GAAATTAATAGTTTAGTAATTCCTTGCGGGGTTTTCATATAATTCATTAATTTTAATAT	20137
OY	143	AGTACCACTTTTAATAAAGGGTTTTATATGAATAAACTTTTAATTTTGTTTGGCAAC	202
Dd	20136	TAAATTTTATGGGAGTAATATCAATATGAAAAAATTTTAACTATGATTTAATTTT	20077
OY	203	TTTTGTGTTTTTTCAGTCTGCTCACCACTATATGTTCTAAAAATNGSCGCTTGGGAT	262
Dd	20076	AGCTTAACAAATACAAATCTTTGCACAACAAGTAAGCTTGA AAAAAGTTGGAGATTT	20017
OY	263	AGTGTGAGAAAAA CTTTTGTGTATGAACTAGACAAGAACTCTAATTTGACCATTT	322
Dd	20016	-----GAAACATTAATGAANAATTAAGGCGAAAAAGCAACATTACTAGCAOCATT	19966
OY	323	TTATTTGAACCTTTTTTTAGGGTTTGGAAATPAGGCTCCTTTGCTCAAGAGATATTCGGA	382
Dd	19965	CTTTTGAAATTAATTTTAACTTTAGGAATFAGATCTTTTGCCAAGGAAATTAATTTG	19906
OY	383	GGTTCTCTAATTTCTGGAATTTGATCGGTTGGTGGTATAGGCGCTTAATACCTGAGGGGCTAT	442
Dd	19905	GGTGGCGCAAGCTCTGGATCTCAGTTATTTAGSAGAAATCTTTGCATAGCTGGAAATAT	19846
OY	443	TTTGATATCAAAAGCGCTTGATGTATTACTA AA AAAAGCTGCTTTCAATGGAATTTGGGT	502
Dd	19845	CTTGCCATPACAGATGATGA A AACAGACCAACACTG-----	19807
OY	503	AAGGAGTTATGTTAGCAGGTGTGTTACTATGCTGTGACAAAGATTAACAGAAATTAAT	562
Dd	19806	CATATAATACAAACGATGGAGTAGCGACGATTAATAGCATCCACATGACCTCACTTAAT	19747
OY	563	CTTCATTTACACTTGTCAATAGTTATATAGGAAGCAATAAAGCTTAATAGCTACT	622
Dd	19746	ATTCATTTTACCTTTTGCAATTAACACAAATGCMAATCTTAA AAAAGACTCGGCATTTGAT	19687
OY	623	TTAGGAGGATTTGAACCTAGTTTGTGATTTGCCAATGGGCCAATCC	667
Dd	19686	ATTGGGGGTTTGAACCAATTTTCATTTTGGAAATAACGGGATTC	19642
RESULT 8			
AC119697/c		173540 bp	DNA linear HTG 18-JUL-2002
LOCUS	Rattus norvegicus clone CH230-135B13,	*** SEQUENCING IN PROGRESS	
DEFINITION	***, 61 unordered pieces.		
ACCESSION	AC119697		
VERSION	AC119697.4	GI:21747171	
KEYWORDS	HTG; HTGS; PHASEI.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Euarlyota, Metazoa; Chordata; Craniata, Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 173540)		
AUTHORS	Munzy D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbida,J., Benton,J., Blimede,K., Blankenburg,K., Bonnin,D., Bouck,J.J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,C., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroil,L., Dedereich,D.A., Delanaye,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Donthuwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,R.K., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,		

REFERENCE JOURNAL	AUTHORS	TITLE	COMMENT	
4202	Corrall, J.H., Nuevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hui, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Maxwell, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokhwo, S., Ogih, M., Okunolu, G., Orangunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Peters, L., Pickens, R., Prims, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, Y., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.	Unpublished	2 (bases 1 to 173540)	
4203	Worley, K.C.	Submitted (01-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4204	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4205	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4206	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4207	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4208	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4209	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4210	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4211	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4212	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4213	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4214	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4215	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4216	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4217	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4218	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4219	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4220	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4221	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4222	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4223	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4224	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4225	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4226	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4227	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4228	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4229	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4230	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4231	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4232	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4233	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4234	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4235	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4236	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4237	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4238	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4239	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4240	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4241	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4242	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4243	Worley, K.C.	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
4244	Worley, K.C.	Submitted (18-JUL-2002		

```

* 76922 77021: gap of unknown length
* 77022 79872: contig of 2851 bp in length
* 79872 79972: gap of unknown length
* 79972 83509: contig of 3537 bp in length
* 83510 83609: gap of unknown length
* 83610 86201: contig of 2592 bp in length
* 86202 86301: gap of unknown length
* 86302 86370: contig of 2269 bp in length
* 86371 88671: gap of unknown length
* 88671 92159: contig of 3489 bp in length
* 92160 92259: gap of unknown length
* 92260 95582: contig of 3323 bp in length
* 95583 95682: gap of unknown length
* 95683 98902: contig of 3220 bp in length
* 98903 99002: gap of unknown length
* 99003 103202: contig of 4200 bp in length
* 103203 103302: gap of unknown length
* 103303 107401: contig of 4099 bp in length
* 107402 107501: gap of unknown length
* 107502 112766: contig of 5265 bp in length
* 112767 112866: gap of unknown length
* 112867 116657: contig of 3791 bp in length
* 11658 116757: gap of unknown length
* 116758 120996: contig of 4239 bp in length
* 120997 121096: gap of unknown length
* 121097 124159: contig of 3063 bp in length
* 124160 124259: gap of unknown length
* 124260 128734: contig of 4475 bp in length

```

Query Match 8.98; Score 67.8; DB 2; Length 173540;  
 Best Local Similarity 45.58; Pred. No. 0.01;  
 Matches 240; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

```

OY 151 TTTTAAATGAAGGGGTTTATGAATAAATCTTAAATTTTGGACACCTTTGTGT 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12741 TGTATATTTCTCTGTGTTTGTATATGCTGTGTTATATTTGTTGCTGTGCTGT 12682
OY 211 TTTTTCAGCTTTCACAACTAATGATCTAAATAATGCGTTGGAGTGAAGTCTGG 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12681 TGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 12622
OY 271 AGAAAACTTTGGTTTGAACAGCAACAGATCTTATGACATTTTATTTGAA 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12621 TGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 12562
OY 331 CCTTTTAAAGGTTTGAATAGGCTCTTGTCTCAAGAGATATCTTGGAGGTTCTCT 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12561 TATGTTCTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 12502
OY 391 TATCTTGGATTTGATGCGGTTGATAGGCTTATACCTTGCAGGAGCTTATTTGGATAT 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12501 TGTGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTAT 12442
OY 451 CAAAGCGCTTATGATGATTAATAAAGCTCTTTCAATGAGACTGGGGTAAGGAGT 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12441 TGTGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 12382
OY 511 TATGTTGAGAGTGGTGTACTATGCGCTGTGACAAAGATTACAGAAATTATCTTCATTT 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12381 TGTGTTGTTTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGC 12322
OY 571 TACATTTGCTAATAGTTAATATAGGAAGCTAAAAATAGCCTTATATGACTTTAGAGAG 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12321 TATGTTGATTTGTTTGTGCTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGC 12262
OY 631 ATTTGAACCTAGTTTGTGATGTGCAATGGCCCAATCCAGTCTCTCTG 677
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12261 TGTGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 12215

```

RESULT 9  
 AE000784/c AE000784 28601 bp DNA linear BCT 17-DEC-1997  
 LOCUS

DEFINITION Borrelia burgdorferi plasmid lp28-3, complete plasmid sequence.  
 ACCESSION AE000784  
 VERSION AE000784.1 GI:2690041  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group  
 1 (bases 1 to 28601)  
 REFERENCE  
 AUTHORS Fraser,C.M., Casjens,S., Huang,M.M., Sutton,G.G., Clayton,R., Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K., Gaim,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D., Salzberg,S., Peterson,J., Kerlavage,A.R., Quackenbush,J., Gocayne,J.D., Weidman,J., Venter,J.C., Palmer,N., Adams,M.D., McDonald,D., Artlich,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K., Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.  
 TITLE  
 JOURNAL Nature 390 (6660), 580-586 (1997)  
 MEDLINE 96065943  
 PUBMED 9403685  
 REFERENCE  
 AUTHORS Fraser,C.M., Casjens,S., Huang,M.M., Sutton,G.G., Clayton,R.A., Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K., Gaim,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D., Salzberg,S., Peterson,J., Kerlavage,A.R., Quackenbush,J., Gocayne,J.D., Weidman,J., Venter,J.C., Palmer,N., Adams,M.D., McDonald,D., Artlich,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K., Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.  
 TITLE  
 JOURNAL Nature 390 (6660), 580-586 (1997)  
 MEDLINE 96065943  
 PUBMED 9403685  
 REFERENCE  
 AUTHORS Fraser,C.M., Casjens,S., Huang,M.M., Sutton,G.G., Clayton,R.A., Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K., Gaim,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D., Salzberg,S., Peterson,J., Kerlavage,A.R., Quackenbush,J., Gocayne,J.D., Weidman,J., Venter,J.C., Palmer,N., Adams,M.D., McDonald,D., Artlich,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K., Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

## FEATURES

## source

1. 28601  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 /plasmid="lp28-3"  
 273..467  
 /gene="BBH01"  
 273..467  
 /gene="BBH01"  
 /note="similar to GP:1655789 percent identity: 100.00;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065980.1"  
 /db\_xref="GI:2690046"  
 /translation="MNRLLFILLVFSNLSNSDNNPLMSNKEKISSEYQINSSN  
 KYSIFKRNSSVKKRYTNHYY"  
 391..858  
 /gene="BBH02"  
 391..858  
 /gene="BBH02"  
 /note="similar to GP:1655790 percent identity: 97.44;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065992.1"  
 /db\_xref="GI:2690048"  
 /translation="MSRQNIQFSNEIQALKDRSTIITNQNDNINSTINYPYQTI  
 LKIEQVDNINQNGTRESGETKLEIINGNISREKDAIOYGSFPAKQVRENQOK  
 ENKDSHIVDDPDKREYIHLIMPISNNADSSSYITNYIINDNLRLISNL"  
 926..1075  
 /gene="BBH03"  
 926..1075  
 /gene="BBH03"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="B. burgdorferi predicted coding region BBH03"

## gene

1. 28601  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 /plasmid="lp28-3"  
 273..467  
 /gene="BBH01"  
 273..467  
 /gene="BBH01"  
 /note="similar to GP:1655789 percent identity: 100.00;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065980.1"  
 /db\_xref="GI:2690046"  
 /translation="MNRLLFILLVFSNLSNSDNNPLMSNKEKISSEYQINSSN  
 KYSIFKRNSSVKKRYTNHYY"  
 391..858  
 /gene="BBH02"  
 391..858  
 /gene="BBH02"  
 /note="similar to GP:1655790 percent identity: 97.44;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065992.1"  
 /db\_xref="GI:2690048"  
 /translation="MSRQNIQFSNEIQALKDRSTIITNQNDNINSTINYPYQTI  
 LKIEQVDNINQNGTRESGETKLEIINGNISREKDAIOYGSFPAKQVRENQOK  
 ENKDSHIVDDPDKREYIHLIMPISNNADSSSYITNYIINDNLRLISNL"  
 926..1075  
 /gene="BBH03"  
 926..1075  
 /gene="BBH03"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="B. burgdorferi predicted coding region BBH03"

## CDS

1. 28601  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 /plasmid="lp28-3"  
 273..467  
 /gene="BBH01"  
 273..467  
 /gene="BBH01"  
 /note="similar to GP:1655789 percent identity: 100.00;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065980.1"  
 /db\_xref="GI:2690046"  
 /translation="MNRLLFILLVFSNLSNSDNNPLMSNKEKISSEYQINSSN  
 KYSIFKRNSSVKKRYTNHYY"  
 391..858  
 /gene="BBH02"  
 391..858  
 /gene="BBH02"  
 /note="similar to GP:1655790 percent identity: 97.44;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="A065992.1"  
 /db\_xref="GI:2690048"  
 /translation="MSRQNIQFSNEIQALKDRSTIITNQNDNINSTINYPYQTI  
 LKIEQVDNINQNGTRESGETKLEIINGNISREKDAIOYGSFPAKQVRENQOK  
 ENKDSHIVDDPDKREYIHLIMPISNNADSSSYITNYIINDNLRLISNL"  
 926..1075  
 /gene="BBH03"  
 926..1075  
 /gene="BBH03"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="B. burgdorferi predicted coding region BBH03"

gene /protein\_id="AAC65997.1"  
/db\_xref="GI:2690053"  
/translation="WKKMKKNCSEIEKTQLEIINNOSEIEKQHLQLEIEFTGVCLLY  
VTHY"  
1045. .1368  
/gene="BBH04"  
1045. .1368  
/note="similar to PID:1142590 GB:AE000793 GB:AE000784  
percent identity: 38.46; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC65986.1"  
/db\_xref="GI:2690042"  
/translation="MPALCDWTLNLLELNYSQKLLKFEYNELKKDNKNCSDPTNNK  
KYLDLEKTKTVKLSFKNQSVMYIKINPLAVKFTSTIDYQYDTDKLKLLELN  
YPTI"  
1498. .1680  
/gene="BBH05"  
1498. .1680  
/note="similar to PID:1142590 GB:AE000793 GB:AE000784  
percent identity: 93.88; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC65991.1"  
/db\_xref="GI:2690047"  
/translation="MNSKNRLSHLLIILYTLNNDLNSKNIGYSGEITRAFTENI  
DKYSKLVEYLFETK"  
complement(2260. .2970)  
/gene="BBH06"  
complement(2260. .2970)  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BBH06"  
/protein\_id="AAC65998.1"  
/db\_xref="GI:2690054"  
/translation="MKKSPFLISYMLISISLSCDVSRLQRFNINELKIFVEKAYYSI  
KIDAYINCTCAVNDIMYISGTFSDQSQKVAQAISEFKDKIVNKFKELEKIIEEK  
PMFLSKLIDDFEILDDQAVNDVDSNARHVADSYKRLRSVLAIESPDYSSFNVS  
KFEVSKRFEVKAQFEVENDLIALCEIVKTIQGMVNDREINSRYNNPYKKEADPL  
GAVALDEATYAIKQTL"  
complement(3083. .3514)  
/gene="BBH07"  
/note="this region contains an authentic frame shift and  
is not the result of a sequencing artifact; similar to  
GB:U03661 PID:458217 percent identity: 63.33; identified  
by sequence similarity; putative; conserved hypothetical  
protein; authentic frameshift"  
complement(3590. .3730)  
/gene="BBH08"  
complement(3590. .3730)  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BBH08"  
/protein\_id="AAC65999.1"  
/db\_xref="GI:2690055"  
/translation="MVGFIKATLTLEIKSPFLSKRSSGCHIGSILTYIAKEDDLAI  
NM"  
complement(3892. .7728)  
/gene="BBH09"  
complement(3892. .7728)  
/gene="BBH09"

[illegible]



QY 544 AAGATTACAGAAATTTCTTCATTTACATTTGCTAATAGTATTAATAGGAAGCTTAA 603  
 Db 27775 CTACTTAATTCATTAATTTTGGCCAGTACTATTGAAGATGATCTCAAAAAATCTTAT 27716

QY 604 AAATGACCTTAATGAGCTTTAGAGAGATTTGAACCTTATTTGATGTCGCAATGGCCCA 663  
 Db 27715 GAATCGAATCATTTGATGAGACTTGGCGGATTTGAACCAATCTTGAATGGAATGAC-- 27658

QY 664 ATTCAGTCTCTTGGGCTTGAACGTCTTCAAAAAAGCATTAATTTTATTTATTTACA 723  
 Db 27657 -----GGATTCACAATGCTCCTTTAAAAAAGTTAATTAATTAATGATATAA 27611

QY 724 AAAA 727  
 Db 27610 GATA 27607

RESULT 10  
 Locus 1592044/c  
 DEFINITION Human DNA sequence from clone RP13-407F1 on chromosome 13, complete sequence.  
 ACCESSION AL592044  
 VERSION AL592044  
 KEYWORDS AL592044.4 GI:16151294  
 SOURCE HTG.  
 ORGANISM human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 236130)  
 AUTHORS Pearce,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquereyes@sanger.ac.uk  
 COMMENT requests: clone@request@sanger.ac.uk  
 On Oct 15, 2001 this sequence version replaced gi:14530013.  
 During sequence assembly data is compared from overlapping clones. While differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPeP; Information on the WormPeP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>  
 RP13-407F1 is from the library RP13-13.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/Dacpac/home.htm>  
 VECTOR: pBACE3.6  
 This sequence is the entire insert of clone RP13-407F1.  
 FEATURES  
 source location/Qualifiers  
 1..236130  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="13"  
 /clone="RP13-407F1"  
 /clone\_id="RP13-13.2"  
 BASE COUNT 79573 a 39654 c 39207 g 77696 t  
 ORIGIN

Query Match 8.2%; Score 62; DB 9; Length 236130;  
 Best Local Similarity 62.6%; Pred. No. 0.09;  
 Matches 114; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 70 AGCAATTAATTTGAAAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTT 129  
 Db 75626 AGAATCATATGAAAAAATATATATCTTTAAATACATGAAAAACCTTAAAAATTTAT 75567

QY 130 AATATGTCACATGATACCACTTTTAATAAGGGGTTTATGATAATTAATTAATTT 189  
 Db 75566 CAATGATTAAGTACTGATGCTTATTAATAAGCAGATTAATTAACATTAATTAAT 75507

QY 190 TGTTTTGGCAACCTT---TGTGTTTTTTCTAGCTTTCGCAACGATTAATGCTTAAAAA 246  
 Db 75506 TGTATGAGATTTTATATCTTGTGATTTTTCACATATGAACTTATGATTAATTTAAAAA 75447

QY 247 TG 248  
 Db 75446 TG 75445

RESULT 11  
 Locus AC117080  
 DEFINITION Dictyostelium discoideum chromosome 2 map 780304-886736 strain AX4, complete sequence.  
 ACCESSION AC117080  
 VERSION AC117080  
 KEYWORDS AC117080.1 GI:20066270  
 SOURCE HTG.  
 ORGANISM Dictyostelium discoideum.  
 Dictyostelium discoideum.  
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.  
 REFERENCE 1 (bases 1 to 106434)  
 AUTHORS Gloeckner,G., Eichinger,L., Szafanski,K., Pachebat,J., Dear,P., Lemmann,R., Baumgart,C., Parra,G., April,J.F., Gulgo,R., Kumpf,K., Tunngal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.  
 TITLE Sequence and Analysis of Chromosome 2 of Dictyostelium  
 JOURNAL Unpublished  
 REMARK The Dictyostelium Genome Sequencing Consortium  
 REFERENCE 2 (bases 1 to 106434)  
 AUTHORS Baumgart,C.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
 COMMENT is available from IMB Jena, Department of Genome Analysis (<http://genome.imb-jena.de/dictyostelium/>) and the University Cologne, Institute for Biochemistry I (<http://www.uni-koeln.de/dictyostelium/project.shtml>)  
 Agency : Deutsche Forschungsgemeinschaft (DFG).  
 FEATURES  
 source location/Qualifiers  
 1..106434  
 /organism="Dictyostelium discoideum"  
 /strain="AX4"  
 /db\_xref="taxon:44689"  
 /chromosome="2"  
 /map="780304-886736"  
 175..1003  
 /note="ORF\_id:dd\_00962"  
 /pseudo  
 /codon\_start=1  
 complement(join(1326..1349,1442..1643,1999..2203,2239..2246))  
 /note="ORF\_id:dd\_00960"  
 /pseudo



[illegible]



RESULT 14  
AC121387/c  
LOCUS  
DEFINITION AC121387 193379 bp DNA linear HTG 22-JUL-2002  
Rattus norvegicus clone CH230-179E15, \*\*\* SEQUENCING IN PROGRESS  
\*\*\* 59 unordered pieces.  
ACCESSION AC121387  
VERSION AC121387.2 GI:21903324  
HTG: HTGS PHASE1.  
KEYWORDS  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 193379)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
Alldbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbarta,J., Benton,J., Blumage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,  
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Niockenkwo,S., Ogih,M., Okunou,G.,  
Orangnye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., RojudoKan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmanl,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 193379)  
Worley,K.C.  
Submitted Submission  
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 193379)  
Worley,K.C.  
Direct Submission  
Submitted (22-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 18, 2002 this sequence version replaced gi:20976316.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GWBS  
Center clone name: CH230-179E15  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 141895 bases at least Q40  
Consensus quality: 148042 bases at least Q30  
Consensus quality: 151839 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 59 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
\* 1183: contig of 1183 bp in length  
\* 1184  
\* 1284: gap of unknown length  
\* 1284  
\* 2773: contig of 1490 bp in length  
\* 2774  
\* 2873: gap of unknown length  
\* 2874  
\* 4113: contig of 1242 bp in length  
\* 4215: gap of unknown length  
\* 4216  
\* 5507: contig of 1292 bp in length  
\* 5508  
\* 5607: gap of unknown length  
\* 5608  
\* 7074: contig of 1467 bp in length  
\* 7075  
\* 7174: gap of unknown length  
\* 7175  
\* 8713: contig of 1359 bp in length  
\* 8714  
\* 8813: gap of unknown length  
\* 8814  
\* 10113: contig of 1300 bp in length  
\* 10114  
\* 10213: gap of unknown length  
\* 10214  
\* 11662: contig of 1449 bp in length  
\* 11663  
\* 11762: gap of unknown length  
\* 11763  
\* 12805: contig of 1043 bp in length  
\* 12806  
\* 12905: gap of unknown length  
\* 12906  
\* 14173: contig of 1268 bp in length  
\* 14174  
\* 14273: gap of unknown length  
\* 14274  
\* 15422: contig of 1149 bp in length  
\* 15423  
\* 15522: gap of unknown length  
\* 15523  
\* 16619: contig of 1097 bp in length  
\* 16620  
\* 16719: gap of unknown length  
\* 16720  
\* 18244: contig of 1525 bp in length  
\* 18245  
\* 18344: gap of unknown length  
\* 18345  
\* 19383: contig of 1039 bp in length  
\* 19384  
\* 19483: gap of unknown length  
\* 19484  
\* 20894: contig of 1411 bp in length  
\* 20895  
\* 20994: gap of unknown length  
\* 20995  
\* 22250: contig of 1256 bp in length  
\* 22251  
\* 22350: gap of unknown length  
\* 22351  
\* 24083: contig of 1733 bp in length  
\* 24084  
\* 24183: gap of unknown length  
\* 24184  
\* 25631: contig of 1448 bp in length  
\* 25632  
\* 25731: gap of unknown length  
\* 25732  
\* 28078: contig of 2347 bp in length  
\* 28178  
\* 28179: gap of unknown length  
\* 28179  
\* 29912: contig of 1734 bp in length  
\* 29913  
\* 30012: gap of unknown length  
\* 30013  
\* 32319: contig of 2307 bp in length  
\* 32320  
\* 32419: gap of unknown length  
\* 32420  
\* 35091: contig of 2672 bp in length  
\* 35092  
\* 35191: gap of unknown length  
\* 35192  
\* 37133: contig of 1942 bp in length  
\* 37134  
\* 37233: gap of unknown length  
\* 37234  
\* 38532: contig of 1299 bp in length  
\* 38533  
\* 38632: gap of unknown length  
\* 38633  
\* 39980: contig of 1348 bp in length  
\* 39981  
\* 40080: gap of unknown length



Db 109157 AAAAAAAAAAAGTTAATATTTTAAAAATATTAAAAAGATTAATAACGATAA 109098  
QY 67 TCAAGCAATATTTGAAAAAGTTAAATTTAAATACTTAAAAACCTTTTAAATTTC 126  
Db 109097 AGGAAAAAAAAAATAAAAAATTTTAAACAACCTCAAAAAAGTGTTTAAATTTT 109038  
QY 127 ATTAAATATGCTACCAATAGTACAGTTTAAATAAGGCTTTTATGATTAATTAAT 186  
Db 109037 ATTATTAATAATATCAATGCCAATCATCTTAAGTATGATTTGATTAATAATTTATTT 108978  
QY 187 TTTTGTTTGGCAACCTTTTGTTTCTTCTAGCTT 222  
Db 108977 TATTTTTTAAATTTTATTTTATTTTATTTT 108942

Search completed: March 27, 2003, 03:05:17  
Job time : 2122.82 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:05:42 ; Search time 167.813 seconds  
(without alignments)  
10185.565 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759  
Sequence: 1 atttgttaagaattgaat.....attctgtattgaatgggtg 759

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N-Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	759	20	AAAX30098
2	596.6	78.6	862	20	AAAX30099
3	558.2	73.5	749	20	AAAX30100
4	540	71.1	540	20	AAAX61505
5	477	62.8	477	20	AAAX61506
6	104.8	13.8	9399	20	AAAX20261
7	92.6	12.2	1761	20	AAAX20310
8	90.6	11.9	7074	20	AAAX20266
9	56.6	7.5	6852	24	ABL70312

10	56.6	7.5	6852	24	AAAX61250	Human gene regulat
11	55.6	7.3	18855	24	ABL332610	Human immune syste
12	55	7.2	53585	20	AAAX20251	Borrelia burgdorfe
13	54.6	7.2	7106	24	ABL33478	Human immune syste
14	54.2	7.1	73334	24	ABL92319	Chemically treated
15	54.2	7.1	73334	24	ABL34125	Human immune syste
16	54	7.1	18624	24	ABL33702	Human immune syste
17	53.8	7.1	396	20	AAAX30115	p11 gene sequence.
18	52.6	6.9	12025	24	ABL33298	Human immune syste
19	52.2	6.9	6012	24	ABL70328	Chemically treated
20	52.2	6.9	6012	24	AAAX61275	Human gene regulat
21	52.2	6.9	6012	24	ABL31371	Signal transductio
22	52.2	6.9	12507	24	ABL32299	Human immune syste
23	52	6.9	74	20	AAAX30103	Borrelia burgdorfe
24	52	6.9	6963	24	ABL32979	Human immune syste
25	51.8	6.8	5586	24	ABK40003	Human chemically p
26	51.6	6.8	56153	22	AAAX46793	Tumour suppressor
27	51.6	6.8	61020	22	AAAX46787	Tumour suppressor
28	50.8	6.7	9084	24	ABL33604	Human immune syste
29	50.6	6.7	13123	24	ABL54364	Chemically treated
30	50.6	6.7	13123	24	ABK31423	Signal transductio
31	50.6	6.7	17674	24	ABL33345	Human immune syste
32	50.4	6.6	18512	24	ABL32976	Human immune syste
33	50.4	6.6	34548	24	ABL70604	Chemically treated
34	50.4	6.6	37515	24	ABO66997	Human angiogenesis
35	50.2	6.6	6381	24	ABL70243	Chemically treated
36	50.2	6.6	6381	24	ABL32966	Human immune syste
37	50.2	6.6	6381	24	ABL34518	Human metastasis a
38	49.8	6.6	8992	24	ABK31421	Signal transductio
39	49.8	6.6	19236	24	ABN80226	Human chemically m
40	49.8	6.6	47108	24	ABK31511	Signal transductio
41	49.6	6.5	5518	24	ABL33331	Human immune syste
42	49.6	6.5	7351	24	ABL32029	Human immune syste
43	49.6	6.5	12356	22	AAAX46510	Tumour suppressor
44	49.4	6.5	7461	24	ABL33784	Human immune syste
45	49.4	6.5	15732	22	AAAX53388	Chemically pretrea

## ALIGNMENTS

RESULT 1					
AAAX30098					
ID	AAAX30098	strand:	DNA;	759	BP.
AC	AAAX30098;				
XX					
DT	17-JUN-1999	(first entry)			
DE	Borrelia burgdorferi	B31 protein encoding cDNA.			
XX					
KW	Borrelia;	P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.			
XX					
OS	Borrelia burgdorferi.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	170..709			
FT		/*tag= a			
XX					
PN	WO9912960-A2.				
XX					
PD	18-MAR-1999.				
XX					
PF	04-SEP-1998;	98WO-IB01424.			
XX					
PR	16-SEP-1997;	97US-0059036.			
PR	10-SEP-1997;	97DK-0001041.			
XX					
PA	(SYMB-)	SYMBICOM AB.			
XX					
PI	Bergstroem S;				
XX					
DR	WPI; 1999-215027/18.				

[illegible]

Dd	601	AAAAAATGACCTTAAATGAGCTTTAGAGAGGATTGAACCTAGTTGGTTTGATGTTCGAATGCG	660
Qy	661	CCAAATCCAGTCTCTTGGGTTTGAACATGCTCTTTCAAAAAAACGTAATTAATTTATTTAT	720
Dy	661	CCAATTCAGTCTCTTGGGTTTGAACATGCTCTTTCAAAAAAACGTAATTAATTTATTTAT	720
Oy	721	ACAAAATGSGTGATTCGCAATTCGTATTGAATGCGCG	759
Dd	721	ACAAAATGSGTGATTCGCAATTCGTATTGAATGCGCG	759
<hr/>			
RESULT 2			
AXX30099	ID	AXX30099 standard; DNA; 862 BP.	
XX	AC	AAx30099;	
XX	DT	17-JUN-1999 (first entry)	
XX	DE	Borrelia afzelii ACAI protein encoding cDNA.	
XX	KW	Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.	
XX	OS	Borrelia afzelii.	
XX	Key	Location/Qualifiers	
FT	CDS	219..755	
FT		/*tag= a	
XX	WM	WO9912360-AZ.	
PD	18-MAR-1999.		
PF	04-SEP-1998;	98WO-IB01424.	
PR	16-SEP-1997;	97US-0059036.	
PR	10-SEP-1997;	97DK-0001041.	
PA	(SYMB-) SYMBICOM AB.		
Pt	Bergstroem S;		
PI	WP1; 1999-215027/18.		
DR	P-PSDB; AAT04279.		
XX	Nucleic acid from Borrelia burgdorferi encoding virulence associated		
PT	protein p13		
PS	Clatm 1; Page 110-111; 118pp; English.		
XX	The present sequence encodes a Borrelia afzelii ACAI protein. The		
CC	present invention describes an isolated nucleic acid (I) that: (1)		
CC	encodes a polypeptide fragment (II) immunologically reactive with		
CC	rabbit polyclonal antibody raised against a 13 kDa polypeptide of		
CC	Borrelia burgdorferi B313, but not reactive with most proteins from		
CC	other spirochaetes; and/or (II) hybridizes under stringent conditions		
CC	to specific nucleic acid sequences but not to genomic DNA from most		
CC	other spirochaetes. (I), (II) and transformed cells are useful in		
CC	vaccines to protect against infection by B. burgdorferi sensu lato.		
CC	(I), (II) and antibodies raised against (II) are used to diagnose such		
CC	infections, in standard immunoassays or amplification/hybridization		
CC	tests. (I) are also used to produce recombinant (II). The 13 kd protein		
CC	is involved in virulence and is highly conserved within B. burgdorferi		
CC	sensu lato, but is absent from other Borrelia species (e.g. those		
CC	responsible for relapsing fever or yaws borreliosis). It should provide		
CC	a more specific and sensitive antibody response, and diagnosis,		
CC	compared with use of whole bacteria as antigen.		
XX	Sequence 862 BP; 265 A; 94 C; 162 G; 341 T; 0 other;		
XX			
Query Match	78.6%; Score 596.6; DB 20; Length 862;		
Best Local Similarity	88.4%; Pred No.2e-112;		
Matches 671; Conservative	0; Mismatches 84; Indels 4; Gaps		



QY	1	ATTCTTAAAGCAATGTGAAATTTGATTTATGTCGAAATCAAGAACTCTATGGGAAG	60
Db	50	ATTATTTAAAGCAATGTGAAATTTGCTAATTTTATGCTCAATCAAGAACTCTATGGGAAG	109
QY	61	CGAATTTCAAGCATAATTTTGAAAAAGTTAAATTTAAATTAACCTTTAAAAACCTTTTGA	120
Db	110	CGAATTTCAAGTAATACCTTTTGAAAAAGTTAAATTTAAATTAAGTTTAAAAACCTTTTGA	169
QY	121	AATTTCTAATATATGTCACCATAGTACAGTTTAAATTAAGGGGTTTTATGATTAACCT	180
Db	170	AATTTCTAATATATGTTACTATATATACAGATTTTAAATTAAGAGGTTTTATGATTAATTT	229
QY	181	TTTATATTTTGTGTTTGGCAACCTTTTGTGTTTTTCTAGCTTTGGTCAAGCTAATGATTC	240
Db	230	TTTATATTTTGTGTTTGGTACCTTTTGTGTTTTTCTAGCTTTGGTCAAGCTAATGATTC	289
QY	241	TAAAAATGATGCTTTGGGATGATGATCTGGAGAAAACTTTTGTTTTATGAACAACGAA	300
Db	290	TAAAA---GCGCTTTTAAATTTGGGACCGGAGAAAAAAGCTTTTATGCTTATGAACAACGAA	346
QY	301	GCAAGATCTATTTGTCACATTTTATTTGAACCTTTTATAGGTTTGGATAGGCTCCCTT	360
Db	347	GAAAGATCTATTTGTCACATTTTATTTGAACCTTTTATAGGTTTGGATAGGCTCCCTT	406
QY	361	TGCTCAAGAGATATTTCTTGAGAGTTCTCTTATTTCTTGATTTGATGCGGTTGGATAG	420
Db	407	TGCTCAAGAGATATTTCTTGAGAGTTTCTCTTATTTCTTGATTTGATGCGGTTGGATAG	466
QY	421	GCTTATATCTTGGCGGGGCTTATTTGGATCAAAAGCGCTTGATGATTAACCTAAAAAGC	480
Db	467	GTTATATCTTACAGAGACTTATTTAGATCAAAAGCTTTGATTAAGATTAAGCTCAAAAGC	526
QY	481	TGCTTTCAATGACCTTGGGGTGAAGGAGTATGTTAGACAGAGTGGTACTATAGGCTGT	540
Db	527	GCTTTTAAAGTGGCTTGGGGTGAAGGAGTATGTTAGACAGAGTGGTACTATAGGCTGT	586
QY	541	GACAGATTAACAGAAATTTTCTGTCATTTACATTTGCTAATAGTATATAGAGAGCT	600
Db	587	GACAGATTAACAGAAATTTTATTTCCGTTTACATTTGCTAATAGTATATAGAGAGACT	646
QY	601	AAAAAATAGCCTTATATGTAGCTTTAGAGAGATTTGAACCTAGTTTGTGATGTGCAATGCG	660
Db	647	GAAAAATAGCCTTATATGTAGCTTTTGAAGGCTTGAAGCTTGTGATTAATATGCGG	706
QY	661	CCAATTCAGTGTCTTGGGTTGAACGTCTTCAAAAAAGCTATTAATTTTA-TTTAT	719
Db	707	CCAAGCTAGGCTCTTGGGTTGAACATCTTTCAAAAAAGCTATTAATTTTAATTTAT	766
QY	720	TACAAAAATGGTATGTCAAATCTCTGATTTGAATGGGT	758
Db	767	TATTAATAATGATGATGCAATTTTGTATTTGCTATTTGCT	805
RESULT 3			
AA30100			
ID AAX30100 standard; DNA: 749 BP.			
XX AAX30100;			
XX 17-JUN-1999 (first entry)			
DE Borrelia garinii IP90 protein encoding cDNA.			
KW Borrelia; p13 antigen; spirochaete; vaccine; infection; diagnosis; ss			
XX Borrelia garinii.			
OS			
XX			
FH Key Location/Qualifiers			
FT CDS 192..725			
FT /+tag- a			
NN W09912960-A2.			

Query Match	Best Local Similarity	Score	DB	Length
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;	73.5%;	558.2;	DB 20;	749;
1	ATTGTTAAAGCAATTGCAATTCATATTTTATGTCGCAATCAAGAACTCTATTGGGAAG	60		
24	ATTGTTAAAGCAATTGCAATTCATATTTTATGTCGCAATCAAGAACTCTATTGGGAAG	83		
61	CGAATTTCAAGCAATTAATTTGAAAAAGCTTAAATTAATAATCAATTAATAAACCTTTTAA	120		
84	CGAATTTGAGACAAATTTGAAAAAGCTTAAATTAATAATCAATTAATAAACCTTTTAA	143		
121	AATTTCAATTAATGCTACCATGACAGTTTAAATAAGGGGTTTTTATGCAATAACT	180		
144	AATTTCAATTAATGCTACCATGACAGTTTAAATAAGGGGTTTTTATGCAATAACT	202		
181	TTTATTTTGTGTTGGCAACCTTTGTTGTTTCTGCTTGTGTCGCAAGCAATGATTC	240		
203	TTTATTTTGTGTTGGCAACCTTTGTTGTTTCTGCTTGTGTCGCAAGCAATGATTC	260		
241	TAAATAAGTGTGTTGGGATGAGTGTGCGAGAAAAACCTTTGTTTATGAAACTAGCAA	300		
261	---TAAAGCACTTTAATCTGGGAGCGGAGAAAAATTTTGGTTTATGAAACTAGCAA	316		
301	GCAAGATCCATATGCTACCAATTTTATGCACTTTTATGAGGTTTGAATAGGCTCTT	360		
317	GAAAGATCTCTGTCACCAATTTTATGCACTTTTATGAGGTTTGAATAGGCTCTT	376		
361	TGCTCAAGAGATATCTGTTGGAGGTTGCTTATCTTGGAATTTGATGGGTTGGATAGG	420		
377	TGCTCAAGAGATATCTGTTGGAGGTTGCTTATCTTGGAATTTGATGGGTTGGATAGG	436		
421	GCTTATACCTTGGGGGCTTATTTGGATATCAAGCGCTTGATGTATTACTATAAAAAAGC	480		



XX New isolated *Borrelia burgdorferi* nucleic acids - used to develop  
PT products for the diagnosis, prevention and treatment of diseases  
PT caused by *Borrelia*, particularly Lyme disease

Claim 1; Page 73; 275pp; English.

This sequence encodes a Borrelia burgdorferi (Bb) protein of the invention, which is suitable for use in a vaccine. The Bb polypeptides can be used in vaccines for eliciting protective antibodies to members of the Borrelia genus, particularly for the use against Lyme disease in humans and animals. They can be used for preventing or attenuating infection caused by a member of the Borrelia genus. The products can also be used for detection of members of the Borrelia genus.

**SQ** Sequence 477 BP; 132 A; 62 C; 114 G; 169 T; 0 other;

Query Match	62.8%	Score 477	DB 20	Length 477
Best Local Similarity	100.0%	Pred. No. 4	1e-88	
Matches 477	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Y	233	AATGATTCCTAAAAATGGTGGCTTTGGGATGAGTCGTGGAGAAAACTTTGGTTATGAA	297
Db	1	AATGATTCCTAAAAATGGTGGCTTTGGGATGAGTCGTGGAGAAAACTTTGGTTATGAA	60
QY	293	ACTAGCAGCAGATTCCTATTGTACCACTTTTATATGAACTTTTATAGGGTTGGAGTA	355
Db	61	ACTAGCAGCAGCAGATTCCTATTGTACCACTTTTATATGAACTTTTATAGGGTTGGAGTA	120
QY	353	GGCCCTCTTGCTCAAGAGATATCTGTGGAGGTCCTCTATCTGTGGATTTGATGGCGTT	412
Db	121	GGCTCTCTTGCTCAAGAGATATCTGTGGAGGTCCTCTATCTGTGGATTTGATGGCGTT	186
QY	413	GGTATAGGGCTTATACTTGGCGGGCTTATTTGGATATCAAGCGCTGTGATGATTACT	472
Db	181	GGTATAGGGCTTATACTTGGCGGGCTTATTTGGATATCAAGCGCTGTGATGATTACT	246
QY	473	AAAAAACTGCTTTTCAATGAGACTTGGGGGTAAAGGAGTATTGTAGCAGGTGTGTACT	533
Db	241	AAAAAACTGCTTTTCAATGAGACTTGGGGGTAAAGGAGTATTGTAGCAGGTGTGTACT	300
QY	533	ATGGCTGTGACAGATTTAACAGAAATATTTCTTCCATTACATTTGCTTAATAGTTAAT	592
Db	301	ATGGCTGTGACAGATTTAACAGAAATATTTCTTCCATTACATTTGCTTAATAGTTAAT	366
QY	593	AGCAAGCTAAAAAATAGCCTTAATGTAGCTTTAGAGAGATTGCAACCTAGTTTGTAGTT	653
Db	361	AGCAAGCTAAAAAATAGCCTTAATGTAGCTTTAGAGAGATTGCAACCTAGTTTGTAGTT	420
Y	653	GCATGGGCCAATCCACTGCTCTGGGTTTGAAGCTGCTTTCAAAAAAGCTATTAA	709
Db	421	GCATGGGCCAATCCACTGCTCTGGGTTTGAAGCTGCTTTCAAAAAAGCTATTAA	477

```
RESULT 6
AAX20261
ID AAX20261 standard; DNA; 9399 BP
```

AC	AAX20261;
XX	
DT	04-MAY-1999 (first-entry)
XX	

Borrellia burgdorferi polynucleotide sequence #14.

KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;  
 KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
 KW Infection; diagnosis; characterisation; detection; ds.  
 ....

*Borrelia burgdorferi*.

PN W09858943-A1.

PD 30-DEC-1998.

XX		
PE	18-JUN-1998;	98MO-US12764.
XX		
PR	03-SEP-1997;	97US-0057483.
PR	20-JUN-1997;	97US-0050355.
PR	22-JUL-1997;	97US-0053344.
PR	22-JUL-1997;	97US-0053377.

PA (HUMA-) HUMAN GENOME SCI INC  
PA (MEDI-) MEDIMMUNE INC.

PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;

XX

DR WPI; 1999-081217/07

PT New isolated *Borrelia burgdorferi* nucleic acids - used to develop  
PT products for the detection, diagnosis, characterisation, prevention  
PT and therapy of infections, particularly Lyme disease

PS Claim 1; Page 925-931; 1128pp; English

CC AAx20248 to AAx20402 represent polynucleotide sequences isolated from  
CC *Borrelia burgdorferi* (Bb). Products derived from Bb can be used for  
CC the detection, diagnosis, characterization, prevention and therapy of  
CC Bb infections, e.g. Lyme disease. They can also be used for the  
CC production of biosynthetic products, e.g. enzymes. *Borrelia* belongs  
CC to a family of motile, spiral-shaped bacteria called spirochetes.  
CC Spirochetes are pathogenic in humans and *Borrelia* causes epidemic  
CC endemic relapsing fever, and Lyme Borreliosis, more commonly known as  
CC Lyme disease.

SQ Sequence 9399 BP; 3470 A; 1281 C; 1049 G; 3595 T; 4 other;

Query Match	13.8%	Score 104.8	DB 20	Length 9399
Best Local Similarity	50.2%	Pred. No. 3.1e-12		
Matches 293	Conservative	0	Mismatches 282	Indels 9
				Gaps 1

[illegible]

OY 614 AATGACCTTTAGAGGATTTGAACCTAGTTTGTGCAAT 657  
 DB 8407 GGCATTGCACCTGGGGCTTGAACCAATTTGACATTGCAAT 8450

RESULT 7  
 AAX20310/c  
 ID AAX20310 standard; DNA; 1761 BP.

XX AAX20310;

XX 04-MAY-1999 (first entry)

XX Borrelia burgdorferi polynucleotide sequence #63.

XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;  
 KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
 infection; diagnosis; characterisation; detection; ds.

XX Borrelia burgdorferi.

XX MO9858943-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12764.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;

XX White OR;

XX WPI; 1999-081217/07.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

XX products for the detection, diagnosis, characterisation, prevention

XX and therapy of infections, particularly Lyme disease

XX Claim 1; Page 1045-1046; 1128bp; English.

XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from

XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for

XX the detection, diagnosis, characterisation, prevention and therapy of

XX Bb infections, e.g. Lyme disease. They can also be used for the

XX production of biosynthetic products, e.g. enzymes. Borrelia belongs

XX to a family of motile, spiral-shaped bacteria called Spirochetes.

XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and

XX endemic relapsing fever, and Lyme borreliosis, more commonly known as

XX Lyme disease.

XX Sequence 1761 BP; 548 A; 267 C; 268 G; 676 T; 2 other;

XX

XX

DB 729 AATATGAAGCGAAAAAGCAACCACTAGACCACTTACTTTGATATATCTTTTGTCTT 670  
 OY 345 TTGAATATAGGCTCTCTGCTCAGAGATATCTTGGAGTCTTATCTTGCAATTG 404  
 DB 669 TAGGATATAGATCTTTTGTTCAGAGATATATTTGGTGGTGCACACTACTTGCTCTC 610  
 OY 405 ATGCGGTTGGTATAGGCTTACTTACCTGGGGGCTTATTGGATATCAAGCGCTTGATG 464  
 DB 609 AAGTCTTGGAA-----GGAATTAATTATATGCTGGACACATGATGATAGAGCTATG 555  
 OY 465 GTATTTACTAAAAAAGCTCTTTTCAATGACCTTGAGGCTTAAAGGATATATGACAGGTG 524  
 DB 554 GAGCGCTTACAGAAAGCACACACAGATATATACCGAGAGCTTATACAGAAATAGAG 495  
 OY 525 TGGTTACTATGCTGTGACAGATTAACAGAAATATCTTCCATTTACATTTGCTATA 584  
 DB 494 GCGTCACAAATTTGATGACATCTTACATTAACGATATATATTCATTTAAATTTGCTATA 435  
 OY 585 GTTATATAGAGAGCTTAAAAATAGCTTATATGATGCTTTAGAGGATTTGAACCTAGT 644  
 DB 434 GATACAAAGCAGACCTTAAAAAAGGCTGCGCATTTGACCTTGGGGGCTTAAACCAATT 375  
 OY 645 TTGATGTTGCAATGGGCCAATCC 667  
 DB 374 TTGACATTGGAATTAACGATTC 352

RESULT 8  
 AAX20266  
 ID AAX20266 standard; DNA; 7074 BP.  
 XX AAX20266;  
 AC AAX20266;  
 XX 04-MAY-1999 (first entry)  
 XX Borrelia burgdorferi polynucleotide sequence #19.  
 DE Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;  
 KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
 infection; diagnosis; characterisation; detection; ds.  
 XX Borrelia burgdorferi.  
 XX MO9858943-A1.  
 XX 30-DEC-1998.  
 XX 18-JUN-1998; 98WO-US12764.  
 XX 03-SEP-1997; 97US-0057483.  
 XX 20-JUN-1997; 97US-0050359.  
 XX 22-JUL-1997; 97US-0053344.  
 XX 22-JUL-1997; 97US-0053377.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX (MEDI-) MEDIMUNE INC.  
 XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;  
 XX White OR;  
 XX WPI; 1999-081217/07.  
 XX New isolated Borrelia burgdorferi nucleic acids - used to develop  
 XX products for the detection, diagnosis, characterisation, prevention  
 XX and therapy of infections, particularly Lyme disease  
 XX Claim 1; Page 951-955; 1128bp; English.  
 XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from  
 XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for  
 XX the detection, diagnosis, characterisation, prevention and therapy of  
 XX Bb infections, e.g. Lyme disease. They can also be used for the  
 XX production of biosynthetic products, e.g. enzymes. Borrelia belongs

CC to a family of motile, spiral-shaped bacteria called Spirochetes.  
 CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and  
 CC endemic relapsing fever, and Lyme borreliosis, more commonly known as  
 CC Lyme disease.

XX Sequence 7074 BP; 3064 A; 906 C; 866 G; 2233 T; 5 other;

Query Match 11.9%; Score 90.6; DB 20; Length 7074;  
 Best Local Similarity 50.4%; Pred. No. 2.3e-09;  
 Matches 325; Conservative 0; Mismatches 289; Indels 31; Gaps 3;

QY 24 TAAATTTATGCTCAATACAGACGCTATTTGGAGCAATTTCAACCAATTAATTTGAA 83  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3016 TATATTTAAATCAAAATAAACCTATTATACAACTCACTTAACCAATTTTCATA 3075  
 QY 84 AAAAGTTAAATTAANA-A-ACCTTAAAAACCTTTTAAATTCATTAATGCTACCAT 142  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3076 GAAATTAATAGCTTAAGTCTCTTGGGGGTTTGGATTAATGCAATTAATTAATTA 3135  
 QY 143 AGTACCACTTTTAAATTAAGGGGTTTATGATTAACCTTAAATTTTGTGGCAAC 202  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3136 TAAATTTATTTGGAGGAGTAATATCAATTAATTAATTTTAACTTATTAATTTT 3195  
 QY 203 TTTTGTGTTTTTCTAGCTTTGCTCAAGCTATGATCTTAAATAATGTCGTTGGAGT 262  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3196 AGCTTAACAATACAAATCTTTGCAACAGATAGCTTGAATAAGTGTGGAAGTAT 3255  
 QY 263 AGTCTGAGAAAACCTTTGGTTTATGAACTACAGCAAGATCTTATGCTACCATTT 322  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3256 -----GAACCATTAATGAAATATTAAGCGAAAAGCAACTATCTAGCACCAATTC 3306  
 QY 323 TTATTAACCTTTTATTTGGGTTTGGAAATAGGCTCTTGGCAAGAGATTTCTTGA 382  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3307 CTTTGAATTTATTTTAACTTTAGCAATAGCAATCTTTGTCAGAGATTTATTTGGT 3366  
 QY 383 GGTCTCTTAATCTTGATTTGATGCGGTTGATATAGGCTTATCTTGGGGGCTTAT 442  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3367 GGTGGCGAGTCTTGATCTCAGTTATTAGAGAGATTAATTTGATAGCTGGAATAT 3426  
 QY 443 TTGATATCAAAAGCGCTGATGATTTACTTAAATAAGCTGCTTTCAATGGACTGGGCT 502  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3427 CTTGGCCATACAGATGAATAACAGCAACCACTG-----GG 3465  
 QY 503 AAGGAGTATGTTAGCAGGTGCTGCTATGCTGTGACAAATTAAGAAATTAAT 562  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3466 CATATAATACACAGATAGGAGTAGGACGATTTATGATCCACATAGCCTCAGCTTAT 3525  
 QY 563 CTTCCATTTACATTTGCTAATAGTTATTAATAGGAAGCTAAATAATAGCTTAATGACT 622  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3526 ATTCATTTACATTTGCAATAACACAAATGCAATCTTAAATAAAGACTCGGCATTTGAT 3585  
 QY 623 TTAGAGAGATTTGAACCTAGTTTGATGTTGCAATGGCCCAATTC 667  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 3586 ATTTGGGGTTTGAACCAATTTGATTTGTAATGAATAGCGGATTC 3630

RESULT 9  
 ABL70312  
 ID ABL70312 standard; DNA; 6852 BP.

XX ABL70312;

XX 01-JUL-2002 (first entry)

XX Chemically treated cell signalling DNA sequence complementary to#101.

XX Cell signalling; cytosine methylation; cell signalling disease;

XX cancer; tumour; cytosolic; ds.

XX Unidentified.

XX WO200202807-A2.

XX

PD 10-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07471.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-154758/20.

PT Nucleic acid, useful for diagnosis and therapy of diseases associated  
 with cell signalling e.g. cancer, comprises chemically modified genomic  
 sequences of genes associated with cell signalling

PS Claim 1; SEQ ID NO 202; 24pp+sequence listing; English.

CC The invention relates to a nucleic acid comprising a sequence of at least  
 CC 18 bases of a segment of chemically pre-treated DNA of genes associated  
 CC with cell signalling. The activity of the modified sequences of the  
 CC invention may be described as cytostatic. The object of the invention is  
 CC to provide the chemically modified DNA of genes associated with cell  
 CC signalling, as well as oligonucleotides and/or PNA-oligonucleotides for  
 CC detecting cytosine methylations, as well as a method which is  
 CC particularly suitable for the diagnosis and/or therapy of genetic and  
 CC epigenetic parameters of genes associated with cell signalling. The  
 CC chemically modified DNA provided by the invention is useful for diagnosis  
 CC and therapy of diseases such as solid tumours and cancer. The sequences  
 CC given in records ABL70111-ABL70626 represent chemically pre-treated  
 CC genomic DNA's of genes associated with cell signalling.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification, but is based on sequence information supplied by the  
 CC European Patent Office.

XX Sequence 6852 BP; 1871 A; 177 C; 1611 G; 3193 T; 0 other;

Query Match 7.5%; Score 56.6; DB 24; Length 6852;  
 Best Local Similarity 49.8%; Pred. No. 0.019;  
 Matches 143; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 56 GGAACGAATTTCAAGCAATTAATTTGAAAAAGTAAATTTAAATTAATTTAAACCTT 115  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 926 GGAAGAGATTTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 985  
 QY 116 TTTTAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 175  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 986 TTTTAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1045  
 QY 176 AAATTTTAAATTTTGTGTTGGCAACCTTTTGTGTTTTCAGCTTCTCAAGCTAAT 235  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 1046 TATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1105  
 QY 236 GATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 295  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 1106 TTTTGAATTTTGTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1165  
 QY 236 AGCAAGCAAGATCTATGCTACCAATTTTAAATTTTAAATTTTAAATTTTAAATTT 342  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 1166 TGTAAAAAGTAAATATGTTAAAGCTAAAGTTAAATTTTAAATTTTAAATTT 1212

RESULT 10

XX AAS61250  
 ID AAS61250 standard; DNA; 6852 BP.

XX AAS61250;

XX 29-JAN-2002 (first entry)

XX Human gene regulation-associated gene oligonucleotide #205.

XX



Query Match	Best Local Similarity	Score	DB	Length	53585;
Matches 232; Conservative	0;	Mismatches 195;	Indels 30;	Gaps 3	
Sequence 53585 BP; 19235 A; 8081 C; 7001 G; 19266 T; 2 other;					
7.2%;	50.8%;	Score 55;	DB 20;	Length 53585;	
Pred. No. 0.051;					
MM20248 to MM20402 represent polynucleotide sequences isolated from					
Borrelia burgdorferi (Bb). Products derived from Bb can be used for					
the detection, diagnosis, characterisation, prevention and therapy of					
Bb infections, e.g. Lyme disease. They can also be used for the					
production of biosynthetic products, e.g. enzymes. Borrelia belongs					
to a family of motile, spiral-shaped bacteria called Spirochetes.					
Spirochetes are pathogenic in humans and Borrelia causes epidemic and					
endemic relapsing fever, and Lyme borreliosis, more commonly known as					
Lyme disease.					
Sequence 53585 BP; 19235 A; 8081 C; 7001 G; 19266 T; 2 other;					
Query Match	7.2%;	Score 55;	DB 20;	Length 53585;	
Best Local Similarity	50.8%;	Pred. No. 0.051;			
Matches 232; Conservative	0;	Mismatches 195;	Indels 30;	Gaps 3	
Sequence 53585 BP; 19235 A; 8081 C; 7001 G; 19266 T; 2 other;					
272 GAAAAACCTTTGCTTATGCAAACTAGCAAGCAAGACTCTTATGACATTTTATGAAAC 331					
52880 GAAAGCTTTTAAACAAATATGATATAAGAGAAAAAATCCATATAGGCCATTTCTTTAAAT 52821					
332 CTTTTTTTTAGGGTTTGGATATAGGCTCTTTGCTCAAGAGATATTTCTTGAGAGTTCTCTT 391					
52820 TTAATTTTGGCCCTTTGGAAATAGATTCCTTTGTCACAAAGGAGATTAATTTGTTGAGGCTCA 52761					
332 ATTCCTGATTTGATGCGGCTGCTATAGGCTTATTAATCTTGCGGGGCTTATTTGATATTC 451					
52760 GTGCTTGGATTTAAATTTATTTAGAGCAAACTCTTTGGGGAACGTGAATTTCTTAATTCAC 52701					
452 AAAGGCGTTGATGATTAATCTAAAAAGCTGCTTTTCATGATGGACTGCGGGGTAAGGAGATT 511					
52700 CGAGAAACACAA-----TTAACCGGATACATATTAATTAAGGGTAG----- 52660					
512 ATGTGACAGAGTGTTGCTACTATGGCTGACAGAAATTAACAGAAATATTTCTTCATTT 571					
52659 -----AGCAGAGATGGTTTAAACATCTACACAGTT-----CACTTATTAATTCATTT 52611					
572 ACATTTGCTAATAGTTATATAGAAAGCTAAAAAATAGCCTTAATAGTACCTTTAGAGAGA 631					
52610 ACATTCGCAAAATCGGCACACATAAATACTTAAAAAAGACATCGAGCCTGAACCTTGACAGC 52551					
632 TTTGAACCTAGTTTGCATGTGCATATGGGCCAATCCAGTGCCTTGGTTGAACGTCT 691					
52550 TTTTGAGCCCAATTTTGATCTTGGAAATAAAGGATTCCAATTAATCTTTAAAAAAGTTAT 52491					
692 TTCAAAAAAGCTATTATTTATTTATTAATCAAAAAAT 728					
52490 TAAATTTAATATAGATAATTAATTTTATATAATCTAAT 52454					
RESULT 13					
ABL33478					
ID ABL33478 standard; DNA; 7106 BP.					
XX ABL33478;					
XX 26-MAR-2002 (first entry)					
XX Human immune system associated gene SEQ ID NO: 1451.					
XX Human; immune system disease; cytosine methylation; antistimatic;					
KM antiatherosclerotic; antianaemic; cytosolic; noctropic;					
KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;					
KM antirheumatic; antiarthritic; antidysbacteric; antiparasitic;					
KM antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;					
KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;					
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;					









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1065.16 Seconds  
(without alignments)  
11540.380 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759  
Sequence: 1 atgtgtaagaatgaat.....attctgtatgaatg99tg 759

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: EST:\*  
2: em\_estba:\*  
3: em\_esthum:\*  
4: em\_estio:\*  
5: em\_estnu:\*  
6: em\_estcov:\*  
7: em\_estpl:\*  
8: em\_estro:\*  
9: em\_hlc:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_est3:\*  
14: gb\_est4:\*  
15: gb\_est5:\*  
16: em\_estfun:\*  
17: em\_estom:\*  
18: gb\_gss:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_inv:\*  
21: em\_gss\_pln:\*  
22: em\_gss\_vrt:\*  
23: em\_gss\_fun:\*  
24: em\_gss\_mam:\*  
25: em\_gss\_mus:\*  
26: em\_gss\_other:\*  
27: em\_gss\_pro:\*  
28: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	66.4	8.7	1101	17	CNS0039G
c 2	66.4	8.7	1101	17	CNS001T2
c 3	64.6	8.5	576	17	CNS035N7
c 4	60.6	8.0	1101	17	CNS00E07
c 5	60.6	8.0	1101	17	CNS0100X
c 6	60.4	8.0	1101	17	CNS00E07

c 7	60.2	7.9	1001	17	CNS0155H	AL105023 Drosophila
c 8	58.8	7.7	661	17	CNS020VJ	AL20980V Tetrarodon
c 9	57.8	7.6	1101	17	CNS003B0	AL064078 Drosophila
c 10	57.6	7.6	1101	17	CNS000D1	AL065414 Drosophila
c 11	56.8	7.5	970	17	CNS0182E	AL108800 Drosophila
c 12	56.4	7.4	1038	17	CNS01TL7	AL166660 Tetrarodon
c 13	56.2	7.4	928	17	CNS000DKY	AL071865 Drosophila
c 14	56	7.4	1201	17	CNS016FX	AL106695 Drosophila
c 15	55.8	7.3	1101	17	CNS0106X	AL098595 Drosophila
c 16	55.6	7.3	524	17	CNS01U90	AL167541 Tetrarodon
c 17	55.6	7.3	1203	17	CNS015WU	AL106008 Drosophila
c 18	55.4	7.3	1200	17	CNS016CO	AL106578 Drosophila
c 19	55	7.2	865	17	BH180441	BH180441 016_H_20-
c 20	55	7.2	865	17	CNS07MAN	AL1617393 T3 end of
c 21	55	7.2	1167	17	CNS07360	AL1427102 clone BA0
c 22	55	7.2	1225	17	CNS0161D	AL106171 Drosophila
c 23	54.8	7.2	860	12	BF617324	BF617324 HSVMEC01
c 24	54.8	7.2	952	17	CNS016RS	AL107132 Drosophila
c 25	54.6	7.2	1101	17	CNS001FB	AL060732 Drosophila
c 26	54.4	7.2	1101	17	CNS00LOO	AL068607 Drosophila
c 27	54	7.1	416	17	AZ045571	AZ045571 T234237b
c 28	54	7.1	809	17	B10155	B10155 FAF14-Sp6.1
c 29	53.8	7.1	1038	17	CNS0108N	AL098657 Drosophila
c 30	53.8	7.1	1094	17	CNS012F2	AL101513 Drosophila
c 31	53.8	7.1	1101	17	CNS00HX9	AL073856 Drosophila
c 32	53.8	7.1	1204	17	CNS016E2	AL106628 Drosophila
c 33	53.6	7.1	862	17	CNS029ER	AL187164 Tetrarodon
c 34	53.6	7.1	868	17	CNS010YV	AL099601 Drosophila
c 35	53.6	7.1	963	17	CNS0075X	AL066857 Drosophila
c 36	53.6	7.1	1184	17	CNS04P4P	AL300850 Tetrarodon
c 37	53.2	7.0	1101	17	CNS00EVL	AL069706 Drosophila
c 38	53	7.0	1101	17	CNS002J2	AL063072 Drosophila
c 39	52.8	7.0	641	10	AV735039	AV735039 AV735039
c 40	52.8	7.0	987	17	CNS014PQ	AL104456 Drosophila
c 41	52.8	7.0	1125	9	AL547503	AL547503 AL547503
c 42	52.6	6.9	996	17	CNS00EFU	AL071063 Drosophila
c 43	52.6	6.9	1201	17	CNS0167M	AL106396 Drosophila
c 44	52.4	6.9	805	17	AB005977	AB005977 Mouse gen
c 45	52.4	6.9	866	17	CNS006MS	AL065764 Drosophila

## ALIGNMENTS

RESULT 1  
CNS0039G/LOCUS 1101 bp DNA linear GSS 03-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL063921.1 GI:4941778  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster.  
Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial



QY 487 TCAATGACGCTTGGGGTAAGGGAGTTATGTTAGACAGTGGTTACTATAGCGGTACAG 546  
Db 856 KTTKKKKKKKKKKGGKKDDAAKKKKKKTKKKKGGKGGKGGKGGKDDDAAK 915  
QY 547 ATTAAACAGAAATTATCTTCATTTACATTTGCTATATATAATAGAACCTAAAAA 606  
Db 916 KKKKKKKKKGGGGKKGGKKGGKKGGKKKKTKKKKKKKKKKKKKKKKKPAADA 975  
QY 607 TAGCCTTATAGTACGTTTAGGAGGATTTGAACCTGTTTGATGTTGCAATGGGCCAATC 666  
Db 976 AAADTTTAAATTTTTKKKKKKDKKTKTKKDDAAAAATKKKKKKKKKKKKKKKKKKRD 1035  
QY 667 CAGTGCCTTCGGGTGTGAGCTGTCTTTTAAAAAAAAGCTATTAAATTTATTAATCAAA 726  
Db 1036 ADKKKKTTGGKDKGGKGGKGGKKKKKTTKTKKKKKKKKKKKKKTTTADKKKKKKKAAD 1095  
QY 727 ATGGGT 732  
Db 1096 DKKKK 1101

ISOLATE	3
LOCUS	CNS035N7
DEFINITION	Tetradon nigroviridis genome survey sequence PUC-ori end of clone 21A06 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL228940.1
VERSION	GI:7887933
KEYWORDS	GSS; genome survey sequence. Tetradon nigroviridis. Tetradon nigroviridis.
SOURCE ORGANISM	Tetradon nigroviridis.

REFERENCE	TITLE	JOURNAL	REFERENCE	TITLE	JOURNAL
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	1 (bases 1 to 576)	Ruest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J.	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	Unpublished	2 (bases 1 to 576)
		Ruest-Crolius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	Unpublished	3 (bases 1 to 576)
		Genoscope.	Direct Submission	Submitted (12-APR-2000)	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/Tetraodon">http://www.genoscope.cns.fr/Tetraodon</a> .

FEATURES	SOURCE	location/Qualifiers
		1..576
		/organism="Tetradodon nigroviridis"
		/db_xref="taxon:99883"
		/clone="214A06"
		/clone_11b="G"
		/note="Genoscope sequence ID : C06G214BA03SP1-end : PUC-Or1"
BASE COUNT	62 a	5 c 18 g 444 t 47 others
ORIGIN		

Query Match	8.5%;	Score 64.6;	DB 17;	Length 576;
Best Local Similarity	42.5%;	Pred. No. 0.012;		
Matches 145;	Conservative 30;	Mismatches 166;	Indels 0;	Gaps 0;

Oy	63	AATTCAGGCATATATTTGGAAAAAGTTAAATTAAATACCTTTAAACACCTTTTAA	122
	64		
	65		
	66		
	67		
	68	AATGAAATTTAAGATTTTATTAATAAANAANNAATTTTTTTTTTTTAAATTTTTTTT	107
	69		
Oy	123	TTTCATTTATATAGTCACCATAGTACAGCTTTTAATAAAGGGCTTTTATGATTAACCTT	182
	124		
	125		
	126		
	127		
	128	TTTTTTTATTTTTTTTTTTTTTAATTTTWTTTTTTTTAAWAAAAGTTTTTTTTTTTTTTT	167
	129		
Oy	183	TAAATTTTGTTTTGGCAACCTTTTGTGTTTTCTAGCTTTTCTCAAGCTAATGATTC	242
	184		
	185		
	186	TAAWTTTWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGTTTATWTTT	227
	187		
	188		
Oy	243	AAATATGTGCGCTTGGGATGAGCTGCTGGAGAAAACCTTTGGTTATGAACATGACAC	302
	244		
	245		
	246		
	247	TTTTTTTTTTTTTTTTTTTTTTTAAWMTTATTTTTTAAWMTTTTTTTTTTKTAAAG	287
	248		
Oy	303	AAGATCCATATGTCACCATTTTTTATGAAACCTTTTTTTAGCGTTTGGAAATAGCGCTT	362
	304		
	305		
	306		
	307		
	308	GGGTTTTTTTWTTTTTTTTTTTTTTMMWMTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	347
	309		
Oy	363	CTCAAGGAGATATTCCTGGAGGCTCTCTATCTCTGATTT	403
	364		
	365		
	366		
	367	TTTTACGATTTWTTAGCACCTTTTTTTTTTTTTTTTTTTT	388
	368		

RESULT 4			
CNS00E07			
LOCUS	1101 bp	DNA	linear
DEFINITION	Drosophila melanogaster genome survey sequence TFR3 end of BAC: BACR2P01 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL069440		
VERSION	AL069440.1	GI:4949583	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Drosophila melanogaster		

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :  
Direct Submission  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage  
BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosagawa and Aaron Mamosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RCRC1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2: cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES	source	Location/Qualifiers
		1. .1101
		/organism="Drosophila melanogaster"
		/db_xref="taxon:7227"
		/clone="BACR23P01"
		/clone_11b="RPCI-98"
		/note="end : TET3"
BASE COUNT	366 a	66 c 104 g 351 t 214 others
ORIGIN		

Query Match	8.0%	Score 60.6	DB 17	Length 1101
Best Local Similarity	35.4%	Pred. No. 0.056		
Matches 109; Conservative	44;	Mismatches 124;	Indels 0;	Gaps 0;

```

OY 4 GTTAAAAAGATTGAATGATATTTTATGTCGAATCAAGAGCTCTATTGGGACGA 63
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 GAAAAATATTTTGAAGATATTAATAAATTAATTAATTAATTAATTAATTA 500
OY 64 ATTCAAGCATATTTTGAAGAAAGTTAAATTAATTAATTAATTAATTAAT 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 501 AATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 560
OY 124 TTGATTAATATGCTACATAGTACGATTTTAAAGGGCTTTTATGATAACTTTT 183
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 561 TTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 620
OY 184 AATTTTGTGTTGGCAACCTTTTGTGTTTCTAGCTTGTCAAGCTAATGATTTCTAA 243
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 621 TATAAATTTTAAATTTATATTAATTAATTAATTAATTAATTAATTAAT 680
OY 244 AATGTGCGTTGGGATGAGTCTGAGAGAAACTT 280
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 681 AATTTAATATTTAAATTAATAAATAAATAAATAAATAAATAAATA 717

RESULT 5
CNS0100X 1101 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACH03G04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL098379.1 GI:5609990
VERSION GSS.
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydriidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source 1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACH03G04"
/clone_1ib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : sp6"

BASE COUNT 195 a 108 c 131 g 161 t 506 others

Query Match 8.0%; Score 60.6; DB 17; Length 1101;
Best Local Similarity 16.8%; Pred. No. 0.056;
Matches 98; Conservative 246; Mismatches 237; Indels 1; Gaps 1;

OY 15 TGAATTTGATTAATTTATGCTCAATCAAGAGCTCTATTGGAGCAATTTCAAGCAA 74
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 KGAATTTGATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 484
OY 75 TAATTTGAAAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 134
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 DGGTTTGTWTKDKTDKTDKTDKTDKTDKTDKTDKTDKTDKTDKTDKTDKTD 544

```

```

OY 135 GCTACCATAGTA-CCAGTTTATAAAGGGCTTTTATGATTAACCTTTAATTTTGT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 545 KKTDRGMRKRTDWTGMDRDRDRTDGRADRAANAKDDADADTDKTTGDMGTTTGA 604
OY 194 TTGGCAACCTTTGTGTTTTTCTGCTTCAAGCTTAATGATTTCAAAAATGTTGG 253
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 605 DTTDDDDAAMWMMKMDMDADDTKAGRKGMGKKKKTKTKTKTKTDTKTGTDRK 664
OY 254 TTTGGATGAGTGTGCAAGAAACCTTTGTTTATGAAGTACAGCAAGACTCTATT 313
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 665 GCKDKDWTDKKRDGKRGDDRDKGTGGDDSKAKDRMANRAKATRAAADAATAAATA 724
OY 314 GTACCATTTTATTAACCTTTTATGAGGTTTGAATGAGCTCTTCTCAAGAGAT 373
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 725 AKATAATATTTTKTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 784
OY 374 ATTCTTGGAGGTTCTTTTCTTTCTTTGATTTGATCGGTTGATAGGCTTTACTTGG 433
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 785 DDADKDKWKGRRDDDRKRWGMRDRAGKKKKKKDKHTAMWADDTGTAMATMTARTATA 844
OY 434 GGGCTTATTGGATATCAAGGCTTGATGATTAATAAAGCTCTTTTCAATGG 493
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 845 KTGKAKATDKBRKRGGGGGRGDRGAGRRKRRKRRKRRKRRKRRKRRKRRKRRK 904
OY 494 ACTTGGGTTAAGGAGTTATGTTAGCAGGTGTTTACTATGCTGTGCAAGATTACA 553
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 905 AGTRKAGDKKADGDMGKRRKMDSKTKMTGTDTTWWGRNRRGRAGARRRR 964
OY 554 GAAATTTATCTTCATTTTACATTTGCTAATAGTATTAATAGG 595
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 965 KGRAGRGAGGKWKWMAADABMTWDMKWKMTDMGK 1006

RESULT 6
CNS00E07 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TE73 end of BAC:
DEFINITION BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069440.1 GI:4949583
VERSION GSS.
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydriidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2: cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source 1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29P01"

```

ORIGIN	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040	2045	2050	2055	2060	2065	2070	2075	2080	2085	2090	2095	2100
1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040	2045	2050	2055	2060	2065	2070	2075	2080	2085	2090	2095	2100	

location/qualifiers  
1. .661  
/organism="Tetradon nigroviridis"





**FEATURES**  
**SOURCE**

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR01J16"
/clone_1kb="RPCR-98"
/note="end : TERT"

```

Query Match	7.6%	Score 57.6:	DB 17	Length 1101,
Best Local Similarity	18.3%	Pred. NC. 0.2:		
Matches 120, Conservative 241;		Mismatches 293;		Indels 2; Gaps 1,

OY	101	ACGTTTAAAAAAGCTTTTTTAATTCATTAAATGCTCCOACAGACCGTTTATAATA	160
Db	426	AACARGCAATGACCGCATWATTCCNNCANNNTWTWDNANNCWTCNTRTSOGSCTKGCCCTC	485
OY	161	GGGGTATTTGAATTAACAATTATATTTTGTGTGGCAACCCTTTTGCTTTTCTAC	220
Db	486	TRTSGCATRBCCATCTTWTNCKGATGCNABTBGMCTMAVMTYYCCCCCTCHK	545
OY	221	TTTGCTCAAGCTAATGATTCATAAAATGCTGCGTTGGATGACTGCTGAGAAAAACT	280
Db	546	THMATKATAKCTAAAYTTTGTATTTWTATTTWKANAAMAAAAMWAAMAAAAAMAGARGT	605
OY	281	TTGGTTTATGAACATGACAGCACAGATCCATATTCACATTTTATATGAACCTTTTATA	340
Db	606	NDDBAKDWKAAADAKKAAAMHNGTWTKDKTKKKKKRRWGAKRAALDITMTATT	665
OY	341	GCGTTTGAATAGGCTCCTTGTGCTCAAGAGATATTCTTGGAGTTCTCTATTTCTTGA	400
Db	666	WKTAKKNAGAANKKADRTKAMKGAKGAGATAKARAARKRWKDCKTKKTKYTTA	725
OY	401	TTTGATGCGGTGGATAGGGGCTTTACTGTGGGGGCGTTATTTGGATATCAAAGCGCT	460
Db	726	TKKTTTTRKRAATBKWKRKKKATAGAMDMDIARRGGRGWDDKRRKRGARDCKDTGK	785
OY	461	GATGGTATTTACTAAAAAAGCTCTTTCAATGACCTTGGGTAAGSAGATTATGTTAGCA	520
Db	786	KKTGDKRDKGGDGDRGGAGGCGDRDRAARARAWPDKDT - BGANDKADKKKAGRK	843
OY	521	GCTGCGTTACTATTCGCTGTGACACAAGATTAAACAGAAATATTCTTCOCATTTACATTGCT	580
Db	844	KMWPRKTKWKPRTDAKKAAMAWMRNRBRKDRGRDRRRRGADKKMKMDDDTAKCD	903
OY	581	AATAGTTTAAATAGAACCTAAAAAATAGCCCTAATGATGCTTAGAGAGATTGCAACT	640
Db	904	WTKMDTWTAGATGAWDDDKAKDGMKNGTDWRTKRMKMGGRGDKGRKRTAGARGDOK	963
OY	641	AGTTTGAATGTTGCATGGCCCAATCCAGTGTCTTGGGTTTGAACGTCTTTCAAAAA	700
Db	964	DKKKGKDDRFDDAKATGDKRTWRPTJTJAKRRRRRRRRRAGRGDRDKDKKRWBARBR	1023
OY	701	AGCTATTATTTTATTTATTTACAAAATGGGCGATTGCAATTCGTATGAATGG	756
Db	1024	GKKKKTKDXDKDDOCKRMGRAKAKADAARKDADAADAAAAMAKKAKDKDAKKK	1079

LOCUS	CNS0182E	970 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN37A19 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108800				
VERSION	ALI08800.1	GI:5629104			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epiphytoidae; Drosophilidae; Drosophila. 1 (bases 1 to 970)				
REFERENCE	Genoscope. Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (Edgp) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.				
FEATURES	location/Qualifiers				
source	1..970 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACN37A19" /clone_1Id="DrosBAC" /plasmid="pBelOBAC11" /note="end : 17"				
BASE COUNT	214 a	147 c	153 g	315 t	141 others
ORIGIN					
Query Match	7.5%; Score 56.8; DB 17; Length 970;				
Best Local Similarity	33.8%; Pred. No. 0.3;				
Matches	79; Conservative 60; Mismatches 95; Indels 0; Gaps 0;				
OY	60	CGCAATTCGAAGCAATAATTGAAAAAAAGTAAATTTAAATACCTTTATTAACCTTTTT	119		
Db	734	GCTTAACCTCMCCHMMHHVCMCTMAAATTTWTTTAAATTAATTAATWTMTT	793		
OY	120	AAATTCATTTAATGCACCATAGTACCAGTTTAAATTAAGAAGGGTTTATGATTAAC	179		
Db	794	YMATTTTWTWTTAAAATTTWAAATWTTTAAAAAMTWMMWWMTTGTTTTATTTTMAWA	853		
OY	180	TTTTAATTTTGTTTGGCAACTTTGTGTTTCTTAGCTTTCCTAACGTAATGATT	239		
Db	854	TTTTTWTTTTWTTTTATTAATATTATTAAATAAATTAATWTWATTWATGATWTMTWT	913		
OY	240	CTAAAAATGGTCGCTTGGATGAGTGTGAGAAAACTTTGGTTTATGAAA	293		
Db	914	WTTTWTTTTAAATWAAATWAAATTAAMWMMWMAAAAWMTTAAWMAWMAAAA	967		
RESULT 12					
CNS01TL/c					
LOCUS	CNS01TL7	1038 bp	DNA	linear	GSS 12-MAY-2000
DEFINITION	tetradon nigriviridis genome survey sequence T7 end of clone 194P6 of library g from Tetradon nigriviridis, genomic survey sequence.				
ACCESSION	AL166660.1	GI:7804398			
VERSION	AL166660.1	GI:7804398			
KEYWORDS	GSS: genome survey sequence.				
SOURCE	Tetradon nigriviridis.				
ORGANISM	Tetradon nigriviridis. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetradon.				





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1671.5 Seconds  
(without alignments)

13040.962 Million cell updates/sec

Title: US-09-508-487-22

Sequence: 749  
1 tagaatttcacaataaa.....attatctgaanaatgggtg 749

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_ph:\*  
7: gb\_pl:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: em\_ba:\*  
15: em\_fun:\*  
16: em\_hum:\*  
17: em\_mu:\*  
18: em\_pl:\*  
19: em\_ov:\*  
20: em\_ph:\*  
21: em\_ro:\*  
22: em\_sts:\*  
23: em\_sy:\*  
24: em\_un:\*  
25: em\_vl:\*  
26: em\_ba:\*  
27: em\_fun:\*  
28: em\_hum:\*  
29: em\_mu:\*  
30: em\_pl:\*  
31: em\_ov:\*  
32: em\_ph:\*  
33: em\_ro:\*  
34: em\_sts:\*  
35: em\_sy:\*  
36: em\_un:\*  
37: em\_vl:\*  
38: em\_ba:\*  
39: em\_fun:\*  
40: em\_hum:\*  
41: em\_mu:\*  
42: em\_pl:\*  
43: em\_ov:\*  
44: em\_ph:\*  
45: em\_ro:\*  
46: em\_sts:\*  
47: em\_sy:\*  
48: em\_un:\*  
49: em\_vl:\*  
50: em\_ba:\*  
51: em\_fun:\*  
52: em\_hum:\*  
53: em\_mu:\*  
54: em\_pl:\*  
55: em\_ov:\*  
56: em\_ph:\*  
57: em\_ro:\*  
58: em\_sts:\*  
59: em\_sy:\*  
60: em\_un:\*  
61: em\_vl:\*  
62: em\_ba:\*  
63: em\_fun:\*  
64: em\_hum:\*  
65: em\_mu:\*  
66: em\_pl:\*  
67: em\_ov:\*  
68: em\_ph:\*  
69: em\_ro:\*  
70: em\_sts:\*  
71: em\_sy:\*  
72: em\_un:\*  
73: em\_vl:\*  
74: em\_ba:\*  
75: em\_fun:\*  
76: em\_hum:\*  
77: em\_mu:\*  
78: em\_pl:\*  
79: em\_ov:\*  
80: em\_ph:\*  
81: em\_ro:\*  
82: em\_sts:\*  
83: em\_sy:\*  
84: em\_un:\*  
85: em\_vl:\*  
86: em\_ba:\*  
87: em\_fun:\*  
88: em\_hum:\*  
89: em\_mu:\*  
90: em\_pl:\*  
91: em\_ov:\*  
92: em\_ph:\*  
93: em\_ro:\*  
94: em\_sts:\*  
95: em\_sy:\*  
96: em\_un:\*  
97: em\_vl:\*  
98: em\_ba:\*  
99: em\_fun:\*  
100: em\_hum:\*  
101: em\_mu:\*  
102: em\_pl:\*  
103: em\_ov:\*  
104: em\_ph:\*  
105: em\_ro:\*  
106: em\_sts:\*  
107: em\_sy:\*  
108: em\_un:\*  
109: em\_vl:\*  
110: em\_ba:\*  
111: em\_fun:\*  
112: em\_hum:\*  
113: em\_mu:\*  
114: em\_pl:\*  
115: em\_ov:\*  
116: em\_ph:\*  
117: em\_ro:\*  
118: em\_sts:\*  
119: em\_sy:\*  
120: em\_un:\*  
121: em\_vl:\*  
122: em\_ba:\*  
123: em\_fun:\*  
124: em\_hum:\*  
125: em\_mu:\*  
126: em\_pl:\*  
127: em\_ov:\*  
128: em\_ph:\*  
129: em\_ro:\*  
130: em\_sts:\*  
131: em\_sy:\*  
132: em\_un:\*  
133: em\_vl:\*  
134: em\_ba:\*  
135: em\_fun:\*  
136: em\_hum:\*  
137: em\_mu:\*  
138: em\_pl:\*  
139: em\_ov:\*  
140: em\_ph:\*  
141: em\_ro:\*  
142: em\_sts:\*  
143: em\_sy:\*  
144: em\_un:\*  
145: em\_vl:\*  
146: em\_ba:\*  
147: em\_fun:\*  
148: em\_hum:\*  
149: em\_mu:\*  
150: em\_pl:\*  
151: em\_ov:\*  
152: em\_ph:\*  
153: em\_ro:\*  
154: em\_sts:\*  
155: em\_sy:\*  
156: em\_un:\*  
157: em\_vl:\*  
158: em\_ba:\*  
159: em\_fun:\*  
160: em\_hum:\*  
161: em\_mu:\*  
162: em\_pl:\*  
163: em\_ov:\*  
164: em\_ph:\*  
165: em\_ro:\*  
166: em\_sts:\*  
167: em\_sy:\*  
168: em\_un:\*  
169: em\_vl:\*  
170: em\_ba:\*  
171: em\_fun:\*  
172: em\_hum:\*  
173: em\_mu:\*  
174: em\_pl:\*  
175: em\_ov:\*  
176: em\_ph:\*  
177: em\_ro:\*  
178: em\_sts:\*  
179: em\_sy:\*  
180: em\_un:\*  
181: em\_vl:\*  
182: em\_ba:\*  
183: em\_fun:\*  
184: em\_hum:\*  
185: em\_mu:\*  
186: em\_pl:\*  
187: em\_ov:\*  
188: em\_ph:\*  
189: em\_ro:\*  
190: em\_sts:\*  
191: em\_sy:\*  
192: em\_un:\*  
193: em\_vl:\*  
194: em\_ba:\*  
195: em\_fun:\*  
196: em\_hum:\*  
197: em\_mu:\*  
198: em\_pl:\*  
199: em\_ov:\*  
200: em\_ph:\*  
201: em\_ro:\*  
202: em\_sts:\*  
203: em\_sy:\*  
204: em\_un:\*  
205: em\_vl:\*  
206: em\_ba:\*  
207: em\_fun:\*  
208: em\_hum:\*  
209: em\_mu:\*  
210: em\_pl:\*  
211: em\_ov:\*  
212: em\_ph:\*  
213: em\_ro:\*  
214: em\_sts:\*  
215: em\_sy:\*  
216: em\_un:\*  
217: em\_vl:\*  
218: em\_ba:\*  
219: em\_fun:\*  
220: em\_hum:\*  
221: em\_mu:\*  
222: em\_pl:\*  
223: em\_ov:\*  
224: em\_ph:\*  
225: em\_ro:\*  
226: em\_sts:\*  
227: em\_sy:\*  
228: em\_un:\*  
229: em\_vl:\*  
230: em\_ba:\*  
231: em\_fun:\*  
232: em\_hum:\*  
233: em\_mu:\*  
234: em\_pl:\*  
235: em\_ov:\*  
236: em\_ph:\*  
237: em\_ro:\*  
238: em\_sts:\*  
239: em\_sy:\*  
240: em\_un:\*  
241: em\_vl:\*  
242: em\_ba:\*  
243: em\_fun:\*  
244: em\_hum:\*  
245: em\_mu:\*  
246: em\_pl:\*  
247: em\_ov:\*  
248: em\_ph:\*  
249: em\_ro:\*  
250: em\_sts:\*  
251: em\_sy:\*  
252: em\_un:\*  
253: em\_vl:\*  
254: em\_ba:\*  
255: em\_fun:\*  
256: em\_hum:\*  
257: em\_mu:\*  
258: em\_pl:\*  
259: em\_ov:\*  
260: em\_ph:\*  
261: em\_ro:\*  
262: em\_sts:\*  
263: em\_sy:\*  
264: em\_un:\*  
265: em\_vl:\*  
266: em\_ba:\*  
267: em\_fun:\*  
268: em\_hum:\*  
269: em\_mu:\*  
270: em\_pl:\*  
271: em\_ov:\*  
272: em\_ph:\*  
273: em\_ro:\*  
274: em\_sts:\*  
275: em\_sy:\*  
276: em\_un:\*  
277: em\_vl:\*  
278: em\_ba:\*  
279: em\_fun:\*  
280: em\_hum:\*  
281: em\_mu:\*  
282: em\_pl:\*  
283: em\_ov:\*  
284: em\_ph:\*  
285: em\_ro:\*  
286: em\_sts:\*  
287: em\_sy:\*  
288: em\_un:\*  
289: em\_vl:\*  
290: em\_ba:\*  
291: em\_fun:\*  
292: em\_hum:\*  
293: em\_mu:\*  
294: em\_pl:\*  
295: em\_ov:\*  
296: em\_ph:\*  
297: em\_ro:\*  
298: em\_sts:\*  
299: em\_sy:\*  
300: em\_un:\*  
301: em\_vl:\*  
302: em\_ba:\*  
303: em\_fun:\*  
304: em\_hum:\*  
305: em\_mu:\*  
306: em\_pl:\*  
307: em\_ov:\*  
308: em\_ph:\*  
309: em\_ro:\*  
310: em\_sts:\*  
311: em\_sy:\*  
312: em\_un:\*  
313: em\_vl:\*  
314: em\_ba:\*  
315: em\_fun:\*  
316: em\_hum:\*  
317: em\_mu:\*  
318: em\_pl:\*  
319: em\_ov:\*  
320: em\_ph:\*  
321: em\_ro:\*  
322: em\_sts:\*  
323: em\_sy:\*  
324: em\_un:\*  
325: em\_vl:\*  
326: em\_ba:\*  
327: em\_fun:\*  
328: em\_hum:\*  
329: em\_mu:\*  
330: em\_pl:\*  
331: em\_ov:\*  
332: em\_ph:\*  
333: em\_ro:\*  
334: em\_sts:\*  
335: em\_sy:\*  
336: em\_un:\*  
337: em\_vl:\*  
338: em\_ba:\*  
339: em\_fun:\*  
340: em\_hum:\*  
341: em\_mu:\*  
342: em\_pl:\*  
343: em\_ov:\*  
344: em\_ph:\*  
345: em\_ro:\*  
346: em\_sts:\*  
347: em\_sy:\*  
348: em\_un:\*  
349: em\_vl:\*  
350: em\_ba:\*  
351: em\_fun:\*  
352: em\_hum:\*  
353: em\_mu:\*  
354: em\_pl:\*  
355: em\_ov:\*  
356: em\_ph:\*  
357: em\_ro:\*  
358: em\_sts:\*  
359: em\_sy:\*  
360: em\_un:\*  
361: em\_vl:\*  
362: em\_ba:\*  
363: em\_fun:\*  
364: em\_hum:\*  
365: em\_mu:\*  
366: em\_pl:\*  
367: em\_ov:\*  
368: em\_ph:\*  
369: em\_ro:\*  
370: em\_sts:\*  
371: em\_sy:\*  
372: em\_un:\*  
373: em\_vl:\*  
374: em\_ba:\*  
375: em\_fun:\*  
376: em\_hum:\*  
377: em\_mu:\*  
378: em\_pl:\*  
379: em\_ov:\*  
380: em\_ph:\*  
381: em\_ro:\*  
382: em\_sts:\*  
383: em\_sy:\*  
384: em\_un:\*  
385: em\_vl:\*  
386: em\_ba:\*  
387: em\_fun:\*  
388: em\_hum:\*  
389: em\_mu:\*  
390: em\_pl:\*  
391: em\_ov:\*  
392: em\_ph:\*  
393: em\_ro:\*  
394: em\_sts:\*  
395: em\_sy:\*  
396: em\_un:\*  
397: em\_vl:\*  
398: em\_ba:\*  
399: em\_fun:\*  
400: em\_hum:\*  
401: em\_mu:\*  
402: em\_pl:\*  
403: em\_ov:\*  
404: em\_ph:\*  
405: em\_ro:\*  
406: em\_sts:\*  
407: em\_sy:\*  
408: em\_un:\*  
409: em\_vl:\*  
410: em\_ba:\*  
411: em\_fun:\*  
412: em\_hum:\*  
413: em\_mu:\*  
414: em\_pl:\*  
415: em\_ov:\*  
416: em\_ph:\*  
417: em\_ro:\*  
418: em\_sts:\*  
419: em\_sy:\*  
420: em\_un:\*  
421: em\_vl:\*  
422: em\_ba:\*  
423: em\_fun:\*  
424: em\_hum:\*  
425: em\_mu:\*  
426: em\_pl:\*  
427: em\_ov:\*  
428: em\_ph:\*  
429: em\_ro:\*  
430: em\_sts:\*  
431: em\_sy:\*  
432: em\_un:\*  
433: em\_vl:\*  
434: em\_ba:\*  
435: em\_fun:\*  
436: em\_hum:\*  
437: em\_mu:\*  
438: em\_pl:\*  
439: em\_ov:\*  
440: em\_ph:\*  
441: em\_ro:\*  
442: em\_sts:\*  
443: em\_sy:\*  
444: em\_un:\*  
445: em\_vl:\*  
446: em\_ba:\*  
447: em\_fun:\*  
448: em\_hum:\*  
449: em\_mu:\*  
450: em\_pl:\*  
451: em\_ov:\*  
452: em\_ph:\*  
453: em\_ro:\*  
454: em\_sts:\*  
455: em\_sy:\*  
456: em\_un:\*  
457: em\_vl:\*  
458: em\_ba:\*  
459: em\_fun:\*  
460: em\_hum:\*  
461: em\_mu:\*  
462: em\_pl:\*  
463: em\_ov:\*  
464: em\_ph:\*  
465: em\_ro:\*  
466: em\_sts:\*  
467: em\_sy:\*  
468: em\_un:\*  
469: em\_vl:\*  
470: em\_ba:\*  
471: em\_fun:\*  
472: em\_hum:\*  
473: em\_mu:\*  
474: em\_pl:\*  
475: em\_ov:\*  
476: em\_ph:\*  
477: em\_ro:\*  
478: em\_sts:\*  
479: em\_sy:\*  
480: em\_un:\*  
481: em\_vl:\*  
482: em\_ba:\*  
483: em\_fun:\*  
484: em\_hum:\*  
485: em\_mu:\*  
486: em\_pl:\*  
487: em\_ov:\*  
488: em\_ph:\*  
489: em\_ro:\*  
490: em\_sts:\*  
491: em\_sy:\*  
492: em\_un:\*  
493: em\_vl:\*  
494: em\_ba:\*  
495: em\_fun:\*  
496: em\_hum:\*  
497: em\_mu:\*  
498: em\_pl:\*  
499: em\_ov:\*  
500: em\_ph:\*  
501: em\_ro:\*  
502: em\_sts:\*  
503: em\_sy:\*  
504: em\_un:\*  
505: em\_vl:\*  
506: em\_ba:\*  
507: em\_fun:\*  
508: em\_hum:\*  
509: em\_mu:\*  
510: em\_pl:\*  
511: em\_ov:\*  
512: em\_ph:\*  
513: em\_ro:\*  
514: em\_sts:\*  
515: em\_sy:\*  
516: em\_un:\*  
517: em\_vl:\*  
518: em\_ba:\*  
519: em\_fun:\*  
520: em\_hum:\*  
521: em\_mu:\*  
522: em\_pl:\*  
523: em\_ov:\*  
524: em\_ph:\*  
525: em\_ro:\*  
526: em\_sts:\*  
527: em\_sy:\*  
528: em\_un:\*  
529: em\_vl:\*  
530: em\_ba:\*  
531: em\_fun:\*  
532: em\_hum:\*  
533: em\_mu:\*  
534: em\_pl:\*  
535: em\_ov:\*  
536: em\_ph:\*  
537: em\_ro:\*  
538: em\_sts:\*  
539: em\_sy:\*  
540: em\_un:\*  
541: em\_vl:\*  
542: em\_ba:\*  
543: em\_fun:\*  
544: em\_hum:\*  
545: em\_mu:\*  
546: em\_pl:\*  
547: em\_ov:\*  
548: em\_ph:\*  
549: em\_ro:\*  
550: em\_sts:\*  
551: em\_sy:\*  
552: em\_un:\*  
553: em\_vl:\*  
554: em\_ba:\*  
555: em\_fun:\*  
556: em\_hum:\*  
557: em\_mu:\*  
558: em\_pl:\*  
559: em\_ov:\*  
560: em\_ph:\*  
561: em\_ro:\*  
562: em\_sts:\*  
563: em\_sy:\*  
564: em\_un:\*  
565: em\_vl:\*  
566: em\_ba:\*  
567: em\_fun:\*  
568: em\_hum:\*  
569: em\_mu:\*  
570: em\_pl:\*  
571: em\_ov:\*  
572: em\_ph:\*  
573: em\_ro:\*  
574: em\_sts:\*  
575: em\_sy:\*  
576: em\_un:\*  
577: em\_vl:\*  
578: em\_ba:\*  
579: em\_fun:\*  
580: em\_hum:\*  
581: em\_mu:\*  
582: em\_pl:\*  
583: em\_ov:\*  
584: em\_ph:\*  
585: em\_ro:\*  
586: em\_sts:\*  
587: em\_sy:\*  
588: em\_un:\*  
589: em\_vl:\*  
590: em\_ba:\*  
591: em\_fun:\*  
592: em\_hum:\*  
593: em\_mu:\*  
594: em\_pl:\*  
595: em\_ov:\*  
596: em\_ph:\*  
597: em\_ro:\*  
598: em\_sts:\*  
599: em\_sy:\*  
600: em\_un:\*  
601: em\_vl:\*  
602: em\_ba:\*  
603: em\_fun:\*  
604: em\_hum:\*  
605: em\_mu:\*  
606: em\_pl:\*  
607: em\_ov:\*  
608: em\_ph:\*  
609: em\_ro:\*  
610: em\_sts:\*  
611: em\_sy:\*  
612: em\_un:\*  
613: em\_vl:\*  
614: em\_ba:\*  
615: em\_fun:\*  
616: em\_hum:\*  
617: em\_mu:\*  
618: em\_pl:\*  
619: em\_ov:\*  
620: em\_ph:\*  
621: em\_ro:\*  
622: em\_sts:\*  
623: em\_sy:\*  
624: em\_un:\*  
625: em\_vl:\*  
626: em\_ba:\*  
627: em\_fun:\*  
628: em\_hum:\*  
629: em\_mu:\*  
630: em\_pl:\*  
631: em\_ov:\*  
632: em\_ph:\*  
633: em\_ro:\*  
634: em\_sts:\*  
635: em\_sy:\*  
636: em\_un:\*  
637: em\_vl:\*  
638: em\_ba:\*  
639: em\_fun:\*  
640: em\_hum:\*  
641: em\_mu:\*  
642: em\_pl:\*  
643: em\_ov:\*  
644: em\_ph:\*  
645: em\_ro:\*  
646: em\_sts:\*  
647: em\_sy:\*  
648: em\_un:\*  
649: em\_vl:\*  
650: em\_ba:\*  
651: em\_fun:\*  
652: em\_hum:\*  
653: em\_mu:\*  
654: em\_pl:\*  
655: em\_ov:\*  
656: em\_ph:\*  
657: em\_ro:\*  
658: em\_sts:\*  
659: em\_sy:\*  
660: em\_un:\*  
661: em\_vl:\*  
662: em\_ba:\*  
663: em\_fun:\*  
664: em\_hum:\*  
665: em\_mu:\*  
666: em\_pl:\*  
667: em\_ov:\*  
668: em\_ph:\*  
669: em\_ro:\*  
670: em\_sts:\*  
671: em\_sy:\*  
672: em\_un:\*  
673: em\_vl:\*  
674: em\_ba:\*  
675: em\_fun:\*  
676: em\_hum:\*  
677: em\_mu:\*  
678: em\_pl:\*  
679: em\_ov:\*  
680: em\_ph:\*  
681: em\_ro:\*  
682: em\_sts:\*  
683: em\_sy:\*  
684: em\_un:\*  
685: em\_vl:\*  
686: em\_ba:\*  
687: em\_fun:\*  
688: em\_hum:\*  
689: em\_mu:\*  
690: em\_pl:\*  
691: em\_ov:\*  
692: em\_ph:\*  
693: em\_ro:\*  
694: em\_sts:\*  
695: em\_sy:\*  
696: em\_un:\*  
697: em\_vl:\*  
698: em\_ba:\*  
699: em\_fun:\*  
700: em\_hum:\*  
701: em\_mu:\*  
702: em\_pl:\*  
703: em\_ov:\*  
704: em\_ph:\*  
705: em\_ro:\*  
706: em\_sts:\*  
707: em\_sy:\*  
708: em\_un:\*  
709: em\_vl:\*  
710: em\_ba:\*  
711: em\_fun:\*  
712: em\_hum:\*  
713: em\_mu:\*  
714: em\_pl:\*  
715: em\_ov:\*  
716: em\_ph:\*  
717: em\_ro:\*  
718: em\_sts:\*  
719: em\_sy:\*  
720: em\_un:\*  
721: em\_vl:\*  
722: em\_ba:\*  
723: em\_fun:\*  
724: em\_hum:\*  
725: em\_mu:\*  
726: em\_pl:\*  
727: em\_ov:\*  
728: em\_ph:\*  
729: em\_ro:\*  
730: em\_sts:\*  
731: em\_sy:\*  
732: em\_un:\*  
733: em\_vl:\*  
734: em\_ba:\*  
735: em\_fun:\*  
736: em\_hum:\*  
737: em\_mu:\*  
738: em\_pl:\*  
739: em\_ov:\*  
740: em\_ph:\*  
741: em\_ro:\*  
742: em\_sts:\*  
743: em\_sy:\*  
744: em\_un:\*  
745: em\_vl:\*  
746: em\_ba:\*  
747: em\_fun:\*  
748: em\_hum:\*  
749: em\_mu:\*  
750: em\_pl:\*  
751: em\_ov:\*  
752: em\_ph:\*  
753: em\_ro:\*  
754: em\_sts:\*  
755: em\_sy:\*  
756: em\_un:\*  
757: em\_vl:\*  
758: em\_ba:\*  
759: em\_fun:\*  
760: em\_hum:\*  
761: em\_mu:\*  
762: em\_pl:\*  
763: em\_ov:\*  
764: em\_ph:\*  
765: em\_ro:\*  
766: em\_sts:\*  
767: em\_sy:\*  
768: em\_un:\*  
769: em\_vl:\*  
770: em\_ba:\*  
771: em\_fun:\*  
772: em\_hum:\*  
773: em\_mu:\*  
774: em\_pl:\*  
775: em\_ov:\*  
776: em\_ph:\*  
777: em\_ro:\*  
778: em\_sts:\*  
779: em\_sy:\*  
780: em\_un:\*  
781: em\_vl:\*  
782: em\_ba:\*  
783: em\_fun:\*  
784: em\_hum:\*  
785: em\_mu:\*  
786: em\_pl:\*  
787: em\_ov:\*  
788: em\_ph:\*  
789: em\_ro:\*  
790: em\_sts:\*  
791: em\_sy:\*  
792: em\_un:\*  
793: em\_vl:\*  
794: em\_ba:\*  
795: em\_fun:\*  
796: em\_hum:\*  
797: em\_mu:\*  
798: em\_pl:\*  
799: em\_ov:\*  
800: em\_ph:\*  
801: em\_ro:\*  
802: em\_sts:\*  
803: em\_sy:\*  
804: em\_un:\*  
805: em\_vl:\*  
806: em\_ba:\*  
807: em\_fun:\*  
808: em\_hum:\*  
809: em\_mu:\*  
810: em\_pl:\*  
811: em\_ov:\*  
812: em\_ph:\*  
813: em\_ro:\*  
814: em\_sts:\*  
815: em\_sy:\*  
816: em\_un:\*  
817: em\_vl:\*  
818: em\_ba:\*  
819: em\_fun:\*  
820: em\_hum:\*  
821: em\_mu:\*  
822: em\_pl:\*  
823: em\_ov:\*  
824: em\_ph:\*  
825: em\_ro:\*  
826: em\_sts:\*  
827: em\_sy:\*  
828: em\_un:\*  
829: em\_vl:\*  
830: em\_ba:\*  
831: em\_fun:\*  
832: em\_hum:\*  
833: em\_mu:\*  
834: em\_pl:\*  
835: em\_ov:\*  
836: em\_ph:\*  
837: em\_ro:\*  
838: em\_sts:\*  
839: em\_sy:\*  
840: em\_un:\*  
841: em\_vl:\*  
842: em\_ba:\*  
843: em\_fun:\*  
844: em\_hum:\*  
845: em\_mu:\*  
846: em\_pl:\*  
847: em\_ov:\*  
848: em\_ph:\*  
849: em\_ro:\*  
850: em\_sts:\*  
851: em\_sy:\*  
852: em\_un:\*  
853: em\_vl:\*  
854: em\_ba:\*  
855: em\_fun:\*  
856: em\_hum:\*  
857: em\_mu:\*  
858: em\_pl:\*  
859: em\_ov:\*  
860: em\_ph:\*  
861: em\_ro:\*  
862: em\_sts:\*  
863: em\_sy:\*  
864: em\_un:\*  
865: em\_vl:\*  
866: em\_ba:\*  
867: em\_fun:\*  
868: em\_hum:\*  
869: em\_mu:\*  
870: em\_pl:\*  
871: em\_ov:\*  
872: em\_ph:\*  
873: em\_ro:\*  
874: em\_sts:\*  
875: em\_sy:\*  
876: em\_un:\*  
877: em\_vl:\*  
878: em\_ba:\*  
879: em\_fun:\*  
880: em\_hum:\*  
881: em\_mu:\*  
882: em\_pl:\*  
883: em\_ov:\*  
884: em\_ph:\*  
885: em\_ro:\*  
886: em\_sts:\*  
887: em\_sy:\*  
888: em\_un:\*  
889: em\_vl:\*  
890: em\_ba:\*  
891: em\_fun:\*  
892: em\_hum:\*  
893: em\_mu:\*  
894: em\_pl:\*  
895: em\_ov:\*  
896: em\_ph:\*  
897: em\_ro:\*  
898: em\_sts:\*  
899: em\_sy:\*  
900: em\_un:\*  
901: em\_vl:\*  
902: em\_ba:\*  
903: em\_fun:\*  
904: em\_hum:\*  
905: em\_mu:\*  
906: em\_pl:\*  
907: em\_ov:\*  
908: em\_ph:\*  
909: em\_ro:\*  
910: em\_sts:\*  
911: em\_sy:\*  
912: em\_un:\*  
913: em\_vl:\*  
914: em\_ba:\*  
915: em\_fun:\*  
916: em\_hum:\*  
917: em\_mu:\*  
918: em\_pl:\*  
919: em\_ov:\*  
920: em\_ph:\*  
921: em\_ro:\*  
922: em\_sts:\*  
923: em\_sy:\*  
924: em\_un:\*  
925: em\_vl:\*  
926: em\_ba:\*  
927: em\_fun:\*  
928: em\_hum:\*  
929: em\_mu:\*  
930: em\_pl:\*  
931: em\_ov:\*  
932: em\_ph:\*  
933: em\_ro:\*  
934: em\_sts:\*  
935: em\_sy:\*  
936: em\_un:\*  
937: em\_vl:\*  
938: em\_ba:\*  
939: em\_fun:\*  
940: em\_hum:\*  
941: em\_mu:\*  
942: em\_pl:\*  
943: em\_ov:\*  
944: em\_ph:\*  
945: em\_ro:\*  
946: em\_sts:\*  
947: em\_sy:\*  
948: em\_un:\*  
949: em\_vl:\*  
950: em\_ba:\*  
951: em\_fun:\*  
952: em\_hum:\*  
953: em\_mu:\*  
954: em\_pl:\*  
955: em\_ov:\*  
956: em\_ph:\*  
957: em\_ro:\*  
958: em\_sts:\*  
959: em\_sy:\*  
960: em\_un:\*  
961: em\_vl:\*  
962: em\_ba:\*  
963: em\_fun:\*  
964: em\_hum:\*  
965: em\_mu:\*  
966: em\_pl:\*  
967: em\_ov:\*  
968: em\_ph:\*  
969: em\_ro:\*  
970: em\_sts:\*  
971: em\_sy:\*  
972: em\_un:\*  
973: em\_vl:\*  
974: em\_ba:\*  
975: em\_fun:\*  
976: em\_hum:\*  
977: em\_mu:\*  
978: em\_pl:\*  
979: em\_ov:\*  
980: em\_ph:\*  
981: em\_ro:\*  
982: em\_sts:\*  
983: em\_sy:\*  
984: em\_un:\*  
985: em\_vl:\*  
986: em\_ba:\*  
987: em\_fun:\*  
988: em\_hum:\*  
989: em\_mu:\*  
990: em\_pl:\*  
991: em\_ov:\*  
992: em\_ph:\*  
993: em\_ro:\*  
994: em\_sts:\*  
995: em\_sy:\*  
996: em\_un:\*  
997: em\_vl:\*  
998: em\_ba:\*  
999: em\_fun:\*  
1000: em\_hum:\*  
1001: em\_mu:\*  
1002: em\_pl:\*  
1003: em\_ov:\*  
1004: em\_ph:\*  
1005: em\_ro:\*  
1006: em\_sts:\*  
1007: em\_sy:\*  
1008: em\_un:\*  
1009: em\_vl:\*  
1010: em\_ba:\*  
1011: em\_fun:\*  
1012: em\_hum:\*  
1013: em\_mu:\*  
1014: em\_pl:\*  
1015: em\_ov:\*  
1016: em\_ph:\*  
1017: em\_ro:\*  
1018: em\_sts:\*  
1019: em\_sy:\*  
1020: em\_un:\*  
1021: em\_vl:\*  
1022: em\_ba:\*  
1023: em\_fun:\*  
1024: em\_hum:\*  
1025: em\_mu:\*  
1026: em\_pl:\*  
1027: em\_ov:\*  
1028: em\_ph:\*  
1029: em\_ro:\*  
1030: em\_sts:\*  
1031: em\_sy:\*  
1032: em\_un:\*  
1033: em\_vl:\*  
1034: em\_ba:\*  
1035: em\_fun:\*  
1036: em\_hum:\*  
1037: em\_mu:\*  
1038: em\_pl:\*  
1039: em\_ov:\*  
1040: em\_ph:\*  
1041: em\_ro:\*  
1042: em\_sts:\*  
1043: em\_sy:\*  
1044: em\_un:\*  
1045: em\_vl:\*  
1046: em\_ba:\*  
1047: em\_fun:\*  
1048: em\_hum:\*  
1049: em\_mu:\*  
1050: em\_pl:\*  
1051: em\_ov:\*  
1052: em\_ph:\*  
1053: em\_ro:\*  
1054: em\_sts:\*  
1055: em\_sy:\*  
1056: em\_un:\*  
1057: em\_vl:\*  
1058: em\_ba:\*  
1059: em\_fun:\*  
1060: em\_hum:\*  
1061: em\_mu:\*  
1062: em\_pl:\*  
1063: em\_ov:\*  
1064: em\_ph:\*  
1065: em\_ro:\*  
1066: em\_sts:\*  
1067: em\_sy:\*  
1068: em\_un:\*  
1069: em\_vl:\*  
1070: em\_ba:\*  
1071: em\_fun:\*  
1072: em\_hum:\*  
1073: em\_mu:\*  
1074: em\_pl:\*  
1075: em\_ov:\*  
1076: em\_ph:\*  
1077: em\_ro:\*  
1078: em\_sts:\*  
1079: em\_sy:\*  
1080: em\_un:\*  
1081: em\_vl:\*  
1082: em\_ba:\*  
1083: em\_fun:\*  
1084: em\_hum:\*  
1085: em\_mu:\*  
1086: em\_pl:\*  
1087: em\_ov:\*  
1088: em\_ph:\*  
1089: em\_ro:\*  
1090: em\_sts:\*  
1091: em\_sy:\*  
1092: em\_un:\*  
1093: em\_vl:\*  
1094: em\_ba:\*  
1095: em\_fun:\*  
1096: em\_hum:\*  
1097: em\_mu:\*  
1098: em\_pl:\*  
1099: em\_ov:\*  
1100: em\_ph:\*  
1101: em\_ro:\*  
1102: em\_sts:\*  
1103: em\_sy:\*  
1104: em\_un:\*  
1105: em\_vl:\*  
1106: em\_ba:\*  
1107: em\_fun:\*  
1108: em\_hum:\*  
1109: em\_mu:\*  
1110: em\_pl:\*  
1111: em\_ov:\*  
1112: em\_ph:\*  
1113: em\_ro:\*  
1114: em\_sts:\*  
1115: em\_sy

JOURNAL Infect. Immun. 69 (5), 3323-3334 (2001)  
 MEDLINE 21189251  
 PUBMED 11292755  
 REFERENCE 2 (bases 1 to 678)  
 AUTHORS Noppa, L., Ostberg, I., and Bergstrom, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden  
 FEATURES  
 source  
 1. 678  
 /organism="Borrelia garinii"  
 /strain="1p90"  
 /db\_xref="taxon:29519"  
 /chromosome="large linear"  
 121. . 654  
 /gene="p13"  
 121. . 654  
 /gene="p13"  
 /codon\_start=-1  
 /transl\_table=11  
 /product="membrane protein p13"  
 /protein\_id="AAB28352.1"  
 /db\_xref="GI:4731125"  
 /translation="MNFELFIVIFCAFSFQODSKSTFENIGAGEKFLVETNKKD  
 SLVFLNLFGLGIGSFAGDILIGSLIGFDVAGILITGAYIDIKDPDNNAKKA  
 DFKWTMGKMLAGVVMAYTRLEIYLPPTFANNYNRKLNINIALMGFEESFDIN  
 MGQASALGFSLSRKSY"  
 BASE COUNT 202 a 77 c 140 g 259 t  
 ORIGIN  
 Query Match 90.5%; Score 678; DB 1; Length 678;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-101;  
 Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 ATGGGCCAAGCTAGTCTGTGGTGGACTGCTTTCAGAAAAAAGCTATTAATTTAT 731  
 Db 601 ATGGGCCAAGCTAGTCTGTGGTGGACTGCTTTCAGAAAAAAGCTATTAATTTAT 660  
 QY 732 TTATCTAGAAAATGGGTG 749  
 Db 661 TTATCTAGAAAATGGGTG 678  
 RESULT 2  
 AE001117/c 14102 bp DNA linear BCT 15-DEC-1997  
 LOCUS  
 DEFINITION Borrelia burgdorferi (section 3 of 70) of the complete genome.  
 ACCESSION AE001117 AE000783  
 VERSION AE001117.1 GI:2687907  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Borrelia burgdorferi.  
 Borrelia burgdorferi  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia  
 burgdorferi group.  
 1 (bases 1 to 14102)  
 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,  
 Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,  
 Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D.,  
 Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S.,  
 Hanson, M., van Vugt, R., Palmer, N., Adams, M.D.,  
 Gocayne, J.D., Weidman, J., Uterback, T., Matthey, L., McDonald, L.,  
 Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K.,  
 Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.  
 Direct Submission  
 Submitted (12-DEC-1997) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 2 (bases 1 to 14102)  
 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,  
 Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,  
 Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D.,  
 Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J.,  
 Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D.,  
 Gocayne, J.D., Weidman, J., Uterback, T., Matthey, L., McDonald, L.,  
 Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K.,  
 Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.  
 Direct Submission  
 Submitted (12-DEC-1997) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 FEATURES  
 source  
 1. 14102  
 /organism="Borrelia burgdorferi"  
 /db\_xref="taxon:139"  
 95. . 733  
 /gene="BB0027"  
 95. . 733  
 /gene="BB0027"  
 /note="hypothetical protein; identified by Glimmer;  
 putative  
 /product="B. burgdorferi predicted coding region BB0027"  
 /transl\_table=11  
 /protein\_id="AAC66429.1"  
 /db\_xref="GI:2687920"  
 /translation="MRKTYRILAVLLIGVNIKIAAANIDPHNTSTIGIDISVGI  
 PFIYNDLSKATPIPLVFGGIGALIKYQHILNNALIGLEKRMFEDINHSNIIINPS  
 SVKRIEYVPIPTFSINIFIDIGELFOIPVFNIGFSLNTYGDNRNNTTNLTEDALPT  
 ISFGSGLMNNENYMARCATASMMWMEFEGNSAKMAHFALVSISLVTVNVKL"  
 741. . 1790  
 /gene="BB0028"  
 741. . 1790  
 /gene="BB0028"  
 /note="hypothetical protein; identified by Glimmer;  
 putative  
 /product="B. burgdorferi predicted coding region BB0028"  
 CDS  
 gene

gene  
CDS  
/protein\_id="AAC66428.1"  
/db\_xref="GI:2687919"  
/translation="MKQRYENVFKRRLILNLLIFLLACSSSESIFSQGNLQIKAHBY  
NLGSSSSRGISLVEPTLYIAAMHLFKKENGKIEKIDLSNVEPIINDIVNSGTYLL  
AONKSEELVEPELTKDQWLTKRFLKAYRFLKSGRQGVAEYLLADKNRREKIFD  
LOGSKTPQATQENDKRFQIJSNEENLTGNSLKTQGMNNRTYTHIDQQAKEIPPIK  
TSIRSSSEYLVTGCGIINNLDTRFKVYSTNNVTPIFIDQGEFSSTFAEEFDALL  
IGSNNGFEAFKTKREGCIFALRAPKSVPEPAGVNGSOLSTGLANDIIPVSNNVTIILTO  
GGLKLEKRKLKE"  
complement(1776. .2207)  
/gene="BB0029"  
complement(1776. .2207)  
/note="similar to SP:P17868 GB:X03897 PID:40141  
PID:1303831 GB:AL009126 percent identity: 32.58;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC66423.1"  
/db\_xref="GI:2687914"  
/translation="MLSKIFVADSCNFKIIPLOKFLHNKSLIYVNSKSELETT  
EVALVEVDNDQSEFLNLADRSLSVITNEDIFAPALLDLSQVANDSGRIFDINNIN  
LYFRSKINPLNKLKIKKTYDEDFNKSRIEKTINFTSLFS"  
complement(2197. .2832)  
/gene="BB0030"  
complement(2197. .2832)  
/note="similar to GB:Z36269 PID:1049113 percent identity:  
30.95; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="signal peptidase I (lepb-1)"  
/protein\_id="AAC66422.1"  
/db\_xref="GI:2687913"  
/translation="MAPYLFEDQRLRRQRKRFKKYFLFLLINFEFTKRVQIEMI  
KSNELPITKNASLFVATHITSEFPIPLKNNDLVLEDFELSNPFLLTIKDFEFLN  
KLFKRSAYKSEILAAVODGSYVRGILNVLNKKDPIEFYLNGLNVSYYKLDPFNTDE  
VTKCFLLKNEEFLINDNLVSLNDSRIRGPIINKNAIYSLVLKAVDQKIY"  
complement(2834. .3814)  
/gene="BB0031"  
complement(2834. .3814)  
/note="similar to PID:1652260 percent identity: 35.98;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="signal peptidase I (lepb-2)"  
/protein\_id="AAC66421.1"  
/db\_xref="GI:2687912"  
/translation="MYLRDLKFAISFLVHYEKLTYRKRYKCRKLRAKRGILNF  
LEDYVAALIEVYVINOYVQAKKIPSGSMETLIDIGDFLPDKSTGPELLIGLFTKN  
GFAAGKESIIILFENPEYKSGVFDFIQRLLYMLTSEIDLDDEYGNPNVRFVLR  
GAFAGKAYIFENKRAKAYIKRGEENFLIEDSYWDLVKNFIRKIKIVANEDYGIYGDFA  
MEVALSOLNINISSTPDPSEFDVARDIDFEERIELEYKLSAFMYVYVYKALMRDY  
GIYVYDGYLPIGDRNRNSHDGRFEGVINKKVLGRTLIILPSPRVGFI"  
3885. .5357  
/gene="BB0032"  
3885. .5357  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="B. burgdorferi predicted coding region BB0032"  
/protein\_id="AAC66427.1"  
/db\_xref="GI:2687918"  
/translation="MKDEVDSNVELFEVKLPILGIEPKYVFLITIIILLSISLTL  
IIPFKRPGVAYIKSNILENTYVYIANKYIGRPLNKKYNAAGVRAKMGKTYE  
ORIKHNKPEYKISQINLELVDPBEKTIKQKQKSLMKRKINENTKILPVSFLS  
SLKSHPKYIKKFLKDSIPYLNSTDEMEFDFLNSKAIYSIDQDNSNDEELWNSLKTFE  
DLNENAIETQVNELENDKDKLITLKNAPWVKTLAKTLDNENIOLISKNKINIKLPGFEK  
INSENAIETQVNELENDKSKNISLSTYVNEEFLIQONVTKVEYODFLKENPFWALN

KNEULIREQVLVDENVYLNKFNQDIGNLEAITGISFSAIEVYANWYSKILPTGFKARLPIS  
 QEWELLYQKPEKNKPLNINIEISKVGFNNLNQSSNFENAIFFKEKNYSSENSNFYSLLI  
 TEIRTSYQONNNNLINASTKASFILKNMSSPNIGFRLIVSKE"  
 5369. .5821  
 /gene="BB0033"  
 5369. .5821  
 /gene="BB0033"  
 /note="similar to GP:2073503 percent identity: 47.68;  
 identified by sequence similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="small protein (smpB)"  
 /protein\_id="AAC66420.1"  
 /db\_xref="GI:2687911"  
 /translation="MTNTNLLLENKRAKFNVEFIEKISCGVILKGTVEKSIKKKLSF  
 NNSFASIKKEELMLEENLHVSQYKRGNIENDELAPRLLIKKRELOLQKKFKEGGT  
 LIPISFLUKSIITVEVGICGKKLYIKREILKOKSIKKDLSREIKT"  
 complement(5892. .6431)  
 /gene="BB0034"  
 complement(5892. .6431)  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BB0034"  
 /protein\_id="AAC66426.1"  
 /db\_xref="GI:2687917"  
 /translation="MKKLLIFVLATPCVFFSSPQAQANDSKAGAGMSKGEULVETSK  
 QDPIVPLNLFLFEGIGISFPAQDIIIGSLILGDAVGIGLILAGVLDIKALDGIIR  
 KAAQMTWKGVMGLAGVVTAVTRLEIILPTFPANSYNRKLKNSLNAVAGGEPSPD  
 VAMQSSALGGEFLSFKRSY"  
 complement(6499. .8379)  
 /gene="BB0035"  
 complement(6499. .8379)  
 /gene="BB0035"  
 /note="similar to GP:1405462 percent identity: 31.66;  
 identified by sequence similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="DNA topoisomerase IV (parC)"  
 /protein\_id="AAC66419.1"  
 /db\_xref="GI:2687910"  
 /translation="MDIRITVLDKQNFLOYSSYVIRDRALASVVDGFRVORRIHSLE  
 MHDGNPRKRVANVGNTMKRYHBDGPTSIYEALVNTAKNDLPIEKGNGNGLPTGDPASA  
 SRYIECKRLPLAEVLYLSKETTYEESYDGRNNEPLLPAKIPVILIQSSEGLAVGA  
 AKILPHNFEILNAAVKSSELLGESYDIYDPDEPTGGIYDVNRYADONGVILVRAIETID  
 EKTIVIRELPFEGETESLISIEKAIRKNTKVSINDPFAENVAIELSPRGVASE  
 VIEKLYHTNCOISISVNLILSERYPVVTIKDLIFENAHLOKILKMLELQKSKI  
 LEKFFVYTFEIDPIEKKITVYKLIETIEENLITIISEVLRHKSPSEREVLQKSKI  
 LKPIRKISLPEIDDKNSDKDIKILKKELEKLSINSNTISRGYSINFIDLLAKYSEHOR  
 KTKISLISKRVKKEIATKMKKVYLNLKEGFAGYSLPEFGEITGNASTYDKLIVRENSY  
 VLNKIEKTEPITDKRNCVALYVDINNSKEQIFSIYFRLNDFYVVKRKKIDKPIITKYV  
 YEFLEKQEDVEFDSINLPEFVEFSYTNKIDIVRIEDINPMVRSRSISGRKISSNNLKVK  
 FK"  
 complement(8379. .10178)  
 /gene="BB0036"  
 complement(8379. .10178)  
 /gene="BB0036"  
 /note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018  
 PID:467396 percent identity: 37.98; identified by sequence  
 similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="DNA topoisomerase IV (parE)"  
 /protein\_id="AAC66418.1"  
 /gene="BB0037"  
 complement(8379. .10178)  
 /gene="BB0037"  
 /note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018  
 PID:467396 percent identity: 37.98; identified by sequence  
 similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="DNA topoisomerase IV (parE)"  
 /protein\_id="AAC66418.1"

D6	6623	TAGAAATTTTCTACAAATAAAGCATTGTTAAAGAATTTGAATGTATATTTTAAGCGCA	6564
OY	61	AATCAGAAGACTCTATTTGGTAGCGAATTTGCAGTACAAATTTGAAAAAGCTTAAATTTA	120
D6	6563	AATCAAGAAGCTCTATTTGGGAAGCCAAATTTCAAGCAAATTTGAAAAAGCTTAAATTTA	6504
OY	121	AATATCTCCAAGAGCTTTTTTAAATTTTCATTAATATGCTACCATTAATACCAG-TTTAAAT	179
D6	6503	AATACCTTTAAAACCTTTTTTAAATTTTCATTAATATGCTACCATTAATGCCAGTTTAAAT	6444
OY	180	AAAGGAGTTTTTATGAATTAAGTTTTTAAATTTTATTTTGTAATCTTTTGCCTTTTCT	239
D6	6443	AAAGGAGTTTTTATGAATTAAGTTTTTATTTTGTGTTGGCAACCTTTTGTTGTTTTCT	6384
OY	240	AGTTTGGCTCAAGATGATCT-----AAAAGCACTTTTAATCTGGGAGCGGAGAAAAA	293
D6	6383	AGCTTTGGCTCAAGCAATTAAGATTTCTAAAAATGGTGGTTGGATAGCTGGAGAAAAA	6324
D6	294	TTTTTGGTTTATGAACCTAATTAAGAAAGATTCCTCTTGACATTTTATTTAGACCTTTT	353
D6	6323	CTTTTGGTTTATGAACCTAATTAAGCAAGAAATCTTATTTGACATTTTATTTAGAACCTTTT	6264
OY	354	TTAGGTTTCGGGATAGATGATCTTTTGTGCTCAAGAGATATCTTTGAGAGTCTTATTCCT	413
D6	6263	TTAGGTTTCGGGATAGATGATGCTCTTGTGCTCAAGAGATATCTTTGAGAGTCTTATTCCT	6204
OY	414	GGAATTTGATGCGGTTGATGATAGGTTTAATTAATCTTACAGAGCTTATTTGGACATCAAGAT	473
D6	6203	GGATTTGATGCGGTTGATGATAGGCTTTTACTTGGGGGGCTTATTTGGATATCAAAACG	6144
OY	474	TTTGATTAATTAATGCTCAAAAGCAATCTTATTTAAGTGGACTTGGGGTAAAGGAATGATGTTG	533
D6	6143	CTTGATTTGATTAATTAATGCAAAAGCAATCTTATTTAAGTGGACTTGGGGTAAAGGAATGATGTTG	6084
OY	534	GCAGGTGTTGTTACTATGAGCTGTGACAAAGATTTAGACAAGAAATGTTCTTCATTTACATTT	593
D6	6083	GCAGGTGTTGTTACTATGAGCTGTGACAAAGATTTAGACAAGAAATTAATCTTCATTTACATTT	6024
OY	594	GCTAATTAATTAATTAATGAGAAAGCAATGCTTAATTAATGAGCTTTGGAGAGATTTGAG	653
D6	6023	GCTAATTAATTAATTAATGAGAAAGCAATGCTTAATTAATGAGCTTTGGAGAGATTTGAA	5964
OY	654	CCTAATTTTGAATTAATTAATGAGCAAGCTATGCTCTTGGTTTGGAGCTGCTTCAAA	713
D6	5963	CCTAATTTTGAATTTGATGCAATGGGCCCAATTCAGATGCTCTTGGTTTGAACCTGCTTCAAA	5904
D6	714	AAAAGCTATTAATTTATTTATCTAGAAAAATGGGTG	749
D6	5903	AAAAGCTATTAATTTATTTATTTATTAACAAAAATGGGTG	5868
RESULT 3			
LOCUS	AF085740	777 bp	DNA linear BCT 24-APR-2001
DEFINITION	Borrelia afzelii membrane protein p13 (p13) gene, complete cds.		
ACCESSION	AF085740		
VERSION	AF085740.1	GI:4731122	
KEYWORDS			
SOURCE			
ORGANISM	Borrelia afzelii. Borrelia afzelii. Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group. 1 (bases 1 to 777) Noppa,L., Ostberg,Y., Lavinovich,M. and Bergstrom,S. P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains Infect. Immun. 69 (5), 3323-3334 (2001)		
REFERENCE	JOURNAL MEDLINE PUBMED AUTHORS	2 (bases 1 to 777) Noppa,L., Ostberg,Y. and Bergstrom,S. Direct Submission	
TITLE	JOURNAL	Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea	

FEATURES		University, Umea S-901 87, Sweden	
SOURCE		Location/Qualifiers	
1..777		/organism="Borrelia afzelii"	
		/strain="ACAI"	
		/db_xref="taxon:29518"	
		/chromosome="large linear"	
121..657			
/gene="p13"			
121..657			
/gene="p13"			
/codon_start=1			
/transl_table=1			
/product="membrane protein p13"			
/protein_id="AAD28361.1"			
/db_xref="GI:4731123			
/translation="MNKELIVLLAFVCFSSFAQADSKAFNLGAGEKLLAETSKK DPVPEPLFLFGIGISFPAQDILGFFLLGEPAVIGILIGAYIDIALKNAPK AAFWKMGKMLAGAVTMAVTRLEIIFTFEANSYNRKLKNSLNIAFGFERSFDI NMGQASLGFELSPKRSY"			
BASE COUNT		228 a	84 c 148 g 316 t 1 others
ORIGIN			
Query Match		72.3%	Score 541.4; DB 1; Length 777;
Best Local Similarity		89.6%	Pred. No. 2.9e-79;
Matches 605; Conservative		0;	Mismatches 66; Indels 4; Gaps 2;
QY	73 CTATTGGTACGGAATTCGAGTACAAATTTGAAAAAGTTAAATTAATAGTCCAA	132	
Db	1 CTAATGGAGAGCAATTCAGATATACCTTTGAAAAAGTTAAATTAATAGTTCAAA	60	
QY	133 AGCCTTTTAAATTCATTAATATGACCAATATACAG-TTAAATAGGGGTTTT	191	
Db	61 AACCTTTTAAATTCATTAATATGATCTATTAATACAGTTTAAATTAAGAGGTTTT	120	
QY	192 ATGAATTAAGTTTAAATTTTATTTTGGTAAATCTTTGCTTTTCTAGTTTGGTCAA	251	
Db	121 ATGATTAATTTTAAATTTATTTGTTGGTAAATCTTTGCTTTTCTAGTTTGGTCAA	180	
QY	252 ---GATGATCTTAAAGACACTTTAATCTGGAGCGGAGAAAAATTTGGTTATGAA	308	
Db	181 GCTGATGATCTTAAAGCGCTTTTAAATTTGGAGCGGAGAAAAATTTAGCTTATGAA	240	
QY	309 ACTAATTAAGAAAGATTCCTCTGATACATTTTATTTGAACCTTTTATGGGTGGGATA	368	
Db	241 ACTAGTAAGAAAGATTCCTCTGATACATTTTATTTGAACCTTTTATGGGTGGGATA	300	
QY	369 GGTCTTTTGCACAGAGATATCCTTGAGAGTTCTTATTTCTGATTTGATCGGTT	428	
Db	301 GGTCTTTTGCACAGAGATATCCTTGAGAGTTCTTATTTCTGATTTGATCGGTT	360	
QY	429 GGTATAGGGTTAATCTTACAGAGCTTATTTGGACATCAAGATTTTGAATTAATGCT	488	
Db	361 GGTATAGGGTTAATCTTACAGAGCTTATTTGAATCAAGCTTGAATTAATGCT	420	
QY	488 AAAAAAGCTGATTTAATGAGCACTGGGGTAAAGGAATGATGTGGCAGGTGCTACT	548	
Db	421 CCAAAAGCGCTTTTAAAGTGAAGCTGGGGTAAAGGAATGATGTGGCAGGTGCTACT	480	
QY	549 ATGGGTGAGCAAGATGACAGAAATTTGTTCCATTTACATTTGCAATTAATTAATAC	608	
Db	481 ATGGGTGAGCAAGATGACAGAAATTTTATTTCCGTTTACATTTGCTAATTAATTAAT	540	
QY	609 AGGAAGCTGAAAAAATAGTCTTAATTAATACCTTTGGAGATTTGAGCCTTAGTTGATTT	668	
Db	541 AGGAAGCTGAAAAAATAGCTTAATTAATACCTTTTGGAGGTTTGAGCCTTAGTTGATTT	600	
QY	669 AACATGGGCCAAGCTAGTGTCTGTGGTTTGGAGCTGTCTTCAAAAAAAGCTAATTAATTT	728	
Db	601 AATATGGGCCAAGCTAGGCTCTGTGGTTTGAACATCATCTTCAAAAAAAGTATTAATTT	660	
QY	729 TATTTATCTAGAAA 743		





CDS 116..1009  
 /gene="BBG01"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BBG01"  
 /protein\_id="AAC6073.1"  
 /db\_xref="GI:2690030"  
 /translation="MRSLFLYLTLGMLGMSCLNDLSLSSNKEOKNNNNVKEYSNVO  
 EDGLNDLYSNOEKOSFTKNFGEMKVEDLINPEPIIPESPKANIPINISIVHOK  
 KTKEDDLIPSTNEKEADEAIKYLEENTLONKSPSELREVPVLEKDEVALISVDYD  
 VIEKTNKTSLSMENKNNRDKINKTITLQNNIKINIELFQALIMTDIAENETRSAP  
 FPDIAQRKLESITIKRLESNNRSTYALELSROALSDANSALSLESFAFKRAEPNVR  
 KRIKELIRHAKVLESLNKK"  
 /gene="BBG02"  
 1047..1928  
 /gene="BBG02"  
 1047..1928  
 /gene="BBG02"  
 /note="similar to PIR:A64689 percent identity: 35.78;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAC66053.1"  
 /db\_xref="GI:2690010"  
 /translation="MEINIOSKLNNKNNKLIFISCSLVSTREPDRNFETYSKNR  
 GVIRPGYIKMHILEIONNIALITRLSKTDFRSHAFVYSKISELSIPLGVFPY  
 IOBDEPBEVKENKFNKFRHLVDVKNKFTAEELIGYVAIILXSNIRDEYELI  
 QIDEPKIFVNDSDIFVTIGLGTDLINSLVAVPININICGSFDETLKONS  
 IEKVFKEETNELYYNQTSPFINVSKEVYVTIGSMQTLKSYLTYYKGRMPKEVEH  
 LEVKITHITIGIOKK"  
 /gene="BBG03"  
 2104..2495  
 /note="This region contains an authentic frame shift and  
 is not the result of a sequencing artifact.; similar to  
 GP:2239281 percent identity: 63.64; identified by sequence  
 similarity; putative; conserved hypothetical protein,  
 authentic frameshift"  
 complement(2750..2857)  
 /gene="BBG04"  
 complement(2750..2857)  
 /gene="BBG04"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BBG04"  
 /protein\_id="AAC66074.1"  
 /db\_xref="GI:2690031"  
 /translation="MKQEAISITFDLEIAVHIKTLIDINLILYINGV"  
 complement(2891..4056)  
 /gene="BBG05"  
 /note="This region contains an authentic frame shift and  
 is not the result of a sequencing artifact.; similar to  
 GP:228714 percent identity: 99.12; identified by sequence  
 similarity; putative; transposase-like protein, authentic  
 frameshift"  
 4208..5368  
 /gene="BBG06"  
 4208..5368  
 /gene="BBG06"  
 4208..5368  
 /gene="BBG06"  
 /note="similar to GB:U03641 PID:458216 percent identity:  
 66.76; identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAC66054.1"  
 /db\_xref="GI:2690011"  
 /translation="MKVETPLKLCIITYINPIRSTMINNKKPNCHNLOOKLIY  
 LLSLAVNSKYNKYTOKNILYCENENKRGOPDTLRMONTLYLEKYEKATYNY  
 YKHGJINGTETIYKLNKSKCECIKINQHRREKQDVFARAVNYLIDNRKNSGSD  
 LVCLNNNNIKERKTIQIEKYOVIVFNKCNBSLKEITPLINDINDIELIKIL

gene 5378..5950  
 /gene="BBG07"  
 5378..5950  
 /gene="BBG07"  
 /note="similar to PID:1143398 PID:1174342 PID:1174350  
 PID:2182755 percent identity: 57.84; identified by  
 sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAC66055.1"  
 /db\_xref="GI:2690012"  
 /translation="KQNNAKSTQLVKPIYRCSSNKKDLPIKIKDNQDTYHKKIMDI  
 YKFLNKKRKNRYISLRGLNOSKVEEFLRLRADKFLGITYGKKRIKILPKYQ  
 VGGIEKYSLSKTYMEVREKRSIFCYFRSLPRLIKENVTYTNKALFSMETTLER  
 OYVDFKRYPOKGPFIKWEKSNLNLH"  
 5911..6678  
 /gene="BBG08"  
 5911..6678  
 /gene="BBG08"  
 /note="similar to PID:1174351 percent identity: 39.26;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="stage 0 sporulation protein J (spooJ)"  
 /protein\_id="AAC66056.1"  
 /db\_xref="GI:2690013"  
 /translation="MDRKLAKKPHITAIASIGVGKSTSSIMESTILSKTKVLLY  
 DLDONAVTSFPTODHPRELININYSLSIKHKFKQVVISKNDIFPSYEL  
 AKRSKGNOPKELMIRNAVNYNLEDYDVYIDTPPSSELNALYADKVIIPVLE  
 RMAVENDPLINDIKELENNPKRKKIHIHRSKVEIGVSTELMSILKEVYLNKF  
 IGEVHSEALKRIIDVAIGKENENYKEYLILEKI"  
 6737..7288  
 /gene="BBG09"  
 6737..7288  
 /gene="BBG09"  
 /note="similar to PID:1143400 percent identity: 43.68;  
 identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAC66057.1"  
 /db\_xref="GI:2690014"  
 /translation="MYNKRKVIINDIRVNAVYINTEERDKREVELKLNRIEDD  
 IRNKNITMKITLLEIRNKLILDGKFFEDFIDFIRIARQAVYIKIADLIFEGKLE  
 EDIIEGIDKTLFNLMDKINSKANLIPPLVRLETOACDFYKMPFANYILED  
 FYCKNRQLIKLIEYKKNOKYS"  
 complement(7483..10779)  
 /gene="BBG10"  
 complement(7483..10779)  
 /gene="BBG10"  
 /note="hypothetical protein; identified by Glimmer;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="B. burgdorferi predicted coding region BBG10"  
 /protein\_id="AAC66075.1"  
 /db\_xref="GI:2690032"  
 /translation="MKLDEITIPLSMSINNOKDAIETLIRKIAEOKFSSLDLKS  
 LEKSAKGSNIEKALNKSSONAIKKNFSLANSVNSVSKAGSKSIGVLYKSGKIG  
 SVKLNANKTGAEFQMLAARPIIMAKAIOAGSTISGIFDGAMDLDGFNEVSTF  
 SDMLGNDLGKSLAESNRARGDETLFIRDAIANVTMLSYGATASVEERIRNEGA  
 AGSSEGLERLAIEYSSVESNOVNLDELALIRAGVDIIDIILAEAGLAGEALYKAA  
 SDKIGDGLNKAALSKATSGSGKFNLTAKAEKTLAQAQOOTAKMSEKLDIGKALE  
 PMNIGFEKVOQLMVAITMPLTWTATVITLTKGLFVLAVYGKVMQKMLKALIT  
 TPVFKLIEGIRYIGKLKELATYSSGFUDRFRKAPFPIIKWOKLIDMSKATYTK  
 NLVTFWKAEREKDTSGSEPRDRKFPDPNKTINKMAEDYQRLQDEITRQNDITN  
 KTGAREQALNLEKTIENKQKPKDYSKIFDPLDLENKILIVGERSVNEFNNSV  
 DEVEYONLLEKESREBEITKTLPHDQVSAIOKLDELNEKKAIVEYKGSFETL  
 NESNRQVVALLEKQVNEYKTAIDRSVEAQKALQKEITDLEMETMLLPAKERASAK

MMASKIOAMKFKVDEHKSOFKELNETNNTIKOYAEKADOTTKSLYSDMIDGLNVK  
NAFMKDIAGKFLNKDTGSEIGEEPHNLJGKVNKNVGELEKMTOMYSEMKGLTKTAA  
GAVRPMGEAVALELNGJLDDPVMGLIKOEKRIKAEKREDEDELEKREVELEK  
LEDREDEIKRREKELSELDEDTTEIEFKLOAKSQJSGEFORLHDVOTETKTK  
KODIETLOJLKEARKLEVERNKLSLEPERIKAOVEDKYNADWMWLGPDRLRK  
TKOILDEILKRLAKVSAGSIEIKLHAGAFVSNKPYMKNSEMSSEFQOPELVR  
ITPAPIDENLRLEAKITAEETIKLOKTOSSVNNFVNFNGVDLADKLVMLKSK  
EHLMEFRMAE"  
complement(10776, .11015)  
/gene="BBG11"  
/complement(10776, .11015)  
/gene="BBG11"  
/note="hypothetical protein; identified by Glimmer;  
putative"

Query Match 13.1%; Score 98.4; DB 1; Length 29766;  
Best Local Similarity 49.7%; Pred. No. 3e-07;  
Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;

QY 1 TTGAATTTTCAACAATTAAGATTTGTAAAAAGATTTGAATTCCTAATTTTATGTTA 60  
DB 1910 TGGCATATCAAAAAAATTAATAGCATTAATTTCAATTTAAAAAGATTTATGAAAACTTT 1969  
QY 61 AATCAAGAAGCTCTATTGGTAAGCAATTCAGTAACAAATTTGAAAAAGTTAAATTTA 120  
DB 1970 ACATTAAATCTGCTTAGAAGCAAGAAATTCACACATATCTTTAGAAAAAGATTCA 2029  
QY 121 AATAGTCCAAAAACCTTTTAAATTCATTAATATAGCTACATAATACAGATTAAATA 180  
DB 2030 ACTCCATTAATTAATTAATTAATTTGCTAATTTGTTTATCAAAACTTAATTAATAG 2089  
QY 181 AAGG---GGTTTATGAATTAAGTTTAAATTTTATTTTGGTAATCTTTGCTTTT 237  
DB 2090 GAGGTAAATTTAATGAAAAAATTTTCACATTAATTAATTTTATGTTAACAATGC 2149  
QY 238 CTAGTTTCTCAGATGATTTCTAAAAAGCACTTTAATCTGGAGCGCGAANAATTTT 297  
DB 2150 AATCTTTGCACAGATAACTGTAAGAAAGGTGTGAGT -GAGATATCGCAGCTGTA 2208  
QY 298 TGGTTATTAACATAATAAGAAATCTCTGTACCATTATTTATGAACCTTTTATAG 357  
DB 2209 TGAATATTAAGCGAAAAAGCAACCATCTACACACATTAATCTTTTGT 2268  
QY 358 GGTTCGAGTAGGTTCTTTGCTCAAGGAGATATCTTGGAGGTCTTATCTTGAT 417  
DB 2269 CTTTAGGGATAGATCTTTTGTTCAGAGATTAATATGTTGGTGGCTGACATCTGGCT 2328  
QY 418 TTGATGCGGTTGTATAGGGTTAATCTTACAGAGACTTATTTGACATCAAGATTTTG 477  
DB 2329 CTCGAATGCTTGGAGG-----AATACTTAATATAGCTGACACATGATCTAGAGAT 2383  
QY 478 ATTAATATCTTAAAAAAGCTGATTTTAAGTGCAGCTTGGGGTAAGGAATGATGTCAG 537  
DB 2384 TTGGAGGCGTTACAGAAACACAGACAGATAACCGAGGAGATTAATCAAGATAG 2443  
QY 538 GTGTGTTCTACTAGCTGTGACAAAGATGACAGAAATTTGTTCCATTTACATTTGCTA 597  
DB 2444 GAGGGCTCAAAATGTAGACATCTACATACAGTGAATTAATTTCCATTTAAATTTGCTA 2503  
QY 598 AATAATTAACAGGAAGTGAATAATAGTCTTAATATAGCTTGGAGATTTGAGCTTA 657  
DB 2504 ATAGATACAGACGACCTTAAAAAAGGCTCGCATTTGCACCTTGGGGGTTGAACCCA 2563  
QY 658 GTTTTGAATTTAAT 673  
DB 2564 ATTTTGACATTTGAAT 2579

RESULT 6  
LOCUS AE001584 52971 bp DNA linear BCT 21-JUN-2002  
DEFINITION Borrelia burgdorferi plasmid lp56, complete plasmid sequence.  
ACCESSION AE001584

VERSION AE001584.1 GI:6382393  
KEYWORDS Lyme disease spirochete.  
SOURCE Borrelia burgdorferi  
ORGANISM Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.  
REFERENCE 1 (bases 1 to 52971)  
Casjens,S., Palmer,N., van Vugt,R., Huang,W.M., Stevenson,B., Rosa,P., Lathigra,R., Sutton,G., Peterson,J., Dodson,R.J., Haft,D., Hickey,E., Gwinn,M., White,O. and Fraser,C.M.  
A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi  
Mol. Microbiol. 35 (3), 490-516 (2000)  
JOURNAL 20138354  
MEDLINE 10672174  
PUBMED 2 (bases 1 to 52971)  
REFERENCE 1  
Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Lathigra,R., White,O., Dodson,R., Hickey,E.K., Gwinn,M., Peterson,J., van Vugt,R., Palmer,N., Haft,D., Rosa,P. and Stevenson,B.  
Direct Submission  
Submitted (11-NOV-1999) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA  
AUTHORS location/Qualifiers  
1. 52971  
/organism="Borrelia burgdorferi"  
/db\_xref="taxon:139"  
/plasmid="lp56"  
279. .548  
/gene="BBQ01"  
279. .548  
/gene="BBQ01"  
/note="similar to GB:U03641 PID:458212 percent identity: 67.53; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA07704.1"  
/db\_xref="GI:6382394"  
/translation="MTKTAIILVARKDAEICISGVNKNKOTENYNKLIVELCLLF  
GYLVHREYVGRIGGLYLFEGFLYGVLLDLRIITNFKCN"  
710. .802  
/gene="BBQ02"  
710. .802  
/gene="BBQ02"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA07738.1"  
/db\_xref="GI:6382428"  
/translation="MRKRIGICILVIEKSIIVGLMFKIVIR"  
856. .1407  
/gene="BBQ03"  
856. .1407  
/gene="BBQ03"  
/note="similar to GB:U31615 GB:U31418 PID:520783  
PID:551742 percent identity: 97.81; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="outer membrane protein, putative"  
/protein\_id="AA07705.1"  
/db\_xref="GI:6382395"  
/translation="MKIIVGFIALALLGCLYLPDNOEAVOTEPENSESIDMSDE  
IYBEGFSSIKLYASEHRLVIEIKKLLISLKDNPNDVDPVADYNEYNKRFIDGS  
EBSKDLIKLFGYKNEQNNKFKSEVYMWLISCISELSPDKISGEGSPERYRMRP  
TAYQDLKVKREIERTNKGIL"

gene  
AE001584 52971 bp DNA linear BCT 21-JUN-2002  
DEFINITION Borrelia burgdorferi plasmid lp56, complete plasmid sequence.  
ACCESSION AE001584

gene  
CDS

96% identity to GB:50778 percent identity: 93.50; identified by sequence similarity; putative; protein p23, authentic frameshift"  
2744..3496  
/gene="BBQ05"  
2744..3496  
/note="similar to GP:1553115 percent identity: 53.30;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="antigen, P35, putative"  
/protein\_id="AAFP07706.1"  
/db\_xref="GI:6382396"

gene  
CDS

96% identity to GP:2239281 percent identity: 79.75;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAFP07707.1"  
/db\_xref="GI:6382397"

translation="MQIPAQEKLEKGVGDITVWAKYESKRATILAPLLNIFLSLGIG  
STVGDIYGGGALGSQVLGLIMAGMTDIDGFVESTAVTGTGVSIGLGLTA  
ASYVDIILIPRFANRYNADLRKLGLAQLAEPPNDIGINDSNYRLKNPTEQFTW  
G"  
complement(4249..4986)  
/gene="BBQ07"  
complement(4249..4986)

gene  
CDS

96% identity to PID:1143400 percent identity: 62.50;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF07708.1"  
/db\_xref="GI:6382398"

translation="MGKKYNNRKEEIVLNRSPTSHMKELTINDNQQDLINYNNLKERE  
IKYNLDKDIDIERMKLIYEIKKEELKYKNFSEFOFIKHPIYAKQAOLYLKVVE  
KYLEAISIDINKETGEPRGIOROLENTLVNKENSLSVNNKMSIRPEMNKEFYTF  
CKEDTRACRTILERLFPTFKREYLSDLINERYKKRNKKRKLYIFIYGAYVLGMNM  
MSVFQPQSOSCARFEWGGLSFLLL"  
complement(5069..5830)  
/gene="BBQ08"  
complement(5069..5830)

gene  
CDS

96% identity to GB:D26185 SP:PJ37522 PID:467381  
PID:580306 GB:AL009126 percent identity: 54.17; identified  
by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="plasmid partition protein, putative"  
/protein\_id="AAP07709.1"  
/db\_xref="GI:6382399"

translation="MDRKESNIITITASPKRGVGYKTTLITFSYLKDKSRKVLIDIDD  
PNLSISYSFYKIYNIDKNCNSYSILKAKDVYGCOCINKINDISIIPHILLENSENSEI  
LNYKDLLDENIRNITVNNVEDYILLDDPPNLFLTKSLNWVTYIIIPVOVESVEWA  
SLSLIMOTINDIKDRFNKSFNISTIEENOFLNRNPFKOVVELLKREYSVIKGRHYHY  
NSVKYLINELLEPSKYERYEERTKNTLEINIINH"  
complement(5803..6339)  
/gene="BBQ09"  
complement(5803..6339)

gene  
CDS

96% identity to GB:U03641 PID:458217 percent identity:  
58.62; identified by sequence similarity; putative"  
/codon\_start=1

```

Db 3629 ATCTTGCACAGAGAACTTGAAAAGTGTTGGAGATATGCAACCTATATGAATAT 3688
Oy 306 GAACTAAATAGAAAGATTCCTGTTACCATTTTATGACCTTTTATAGGTTCCGG 365
Db 3689 GAAAGTAAAAAGCAACCATATATACACCATATCTTTGACATCTTTTATCTTAGGA 3748
Oy 366 ATAGTCTCTTGTCTCAGGAGATATCCTTGAGGTTCTCTTATCTTGATTTGATGG 425
Db 3749 ATGAGATCTTTGTTGAGAGATATATGTTGGTGGGACACTGCTGCTCAAGTG 3808
Oy 426 GTTGCTATAGGCTTAATACCTTACAGAGCTTATTTGACATCAAGATTTTATATAT 485
Db 3809 CTTGAGAGAACTATATATATGCTGTGATACATACCGGGGATTTGATTCCTTACAGA 3868
Oy 486 GCTAAAAAGCTGATTTTATAGTGACTTGGGTAAGGAGATGTTGGCAGGTGTGTT 545
Db 3869 AGCAGCAGCAACGATATTAACCGAGAGATATATCAGCAAT-----ACGAGGGCTC 3919
Oy 546 ACTAGGCTGTGACAAAGATGTGACAGAAATGTTCTCCATTTACATTTGCTAATATAT 605
Db 3920 ACAATGACGATCCTACATCCTGCAATATATCATTCATTTAAATTTGCTAATAGATAC 3979
Oy 606 AACAGGAAGCTGAAAAATAGTCTAATATAGCCTTGGAGATTTGAGGCTAGTTTGTAT 665
Db 3980 AACGAGACCTTAAAAAAGACTCGGCATTTGCACCTTGGCGGCTTGAAACCAATTTGAC 4039
Oy 666 ATTAACAT 673
Db 4040 ATTGGAAT 4047

RESULT 7
AE000789/c 27323 bp DNA linear BCT 17-DEC-1997
LOCUS Borrelia burgdorferi plasmid lp28-4, complete plasmid sequence.
DEFINITION AE000789
ACCESSION AE000789
VERSION AE000789.1 GI:2690079
KEYWORDS
SOURCE
ORGANISM Borrelia burgdorferi.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
burgdorferi group.
REFERENCE 1 (bases 1 to 27323)
AUTHORS Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.,
Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,
Gwin,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D.,
Peterson,J., Kerlavage,A.R., Quackenbush,J., Salzberg,S.,
Hanson,M., van Vugt,R., Palmer,N., Adams,M.D., Gocayne,J.,
Venter,J.C. et al.
TITLE Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi
JOURNAL Nature 390 (6660), 580-586 (1997)
MEDLINE 98065943
PUBMED 9403685
REFERENCE 2 (bases 1 to 27323)
AUTHORS Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.A.,
Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,
Gwin,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D.,
Richardson,D., Peterson,J., Kerlavage,A.R., Quackenbush,J.,
Salzberg,S., Hanson,M., van Vugt,R., Palmer,N., Adams,M.D.,
Gocayne,J.D., Weidman,J., Utterback,T., Watthey,L., McDonald,L.,
Artlieb,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K.,
Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1997) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. 27323
/organism="Borrelia burgdorferi"
/db_xref="taxon:139"
/plasmid="lp28-4"
174..608
gene

```

```

CDS
/gene="BB101"
174..608
/feature="BB101"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB101"
/protein_id="AAC6211.1"
/db_xref="GI:2690108"
/translation="MENQYEFNLEFFIKIIFLYLYIMKSYKEQOEQIDHACRILI
LTERFEINLEENSOKTLKRYENLNKNLPPSNSTMKYLYKOLEKEIKIARF
YEKNDOSLIYKNTLEKIMKLELEPKORIQKNTT"
736..1104
/gene="BB102"
736..1104
/note="BB102"
/note="similar to GP:1655791 percent identity: 66.67;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66189.1"
/db_xref="GI:2690086"
/translation="MDGVINDPLVARMKKQIKENKILITLVKTLDHNNKELYSANK
TNYVLIQNNENEAIAKTYQLRVNRTKILELEIEIKNPVILKRPTRKENSFGLY
TLSPLEVCCTKIYMLHPNI"
complement(1847..1972)
/gene="BB103"
complement(1847..1972)
/note="BB103"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB103"
/protein_id="AAC66212.1"
/db_xref="GI:2690109"
/translation="WVYVLDNFSLYFGSLACWDEYVWVWISILSEIVSFL"
complement(2124..2219)
/gene="BB104"
complement(2124..2219)
/note="BB104"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB104"
/protein_id="AAC66213.1"
/db_xref="GI:2690110"
/translation="WGRTIOKFLIKHLGLMALPKIGSKYFTFY"
complement(2188..2310)
/gene="BB105"
complement(2188..2310)
/note="BB105"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB105"
/protein_id="AAC66214.1"
/db_xref="GI:2690111"
/translation="MPYLVNWSNVQKETILKALKLFQDRKSGNNPRTFN"
2536..3351
/gene="BB106"
2536..3351
/feature="BB106"
/note="similar to GB:D26562 SP:P24247 GB:M83735 PID:473817
PID:457107 percent identity: 33.62; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="pfs protein (pfs)"

```

gene  
CDS  
/protein\_id="AAC66190.1"  
/db\_xref="GI:2690087"  
/translation="MKIKINIANIILPELSTLVNSWGMSKININPNENSQSNIIILISAT  
KAEIEILINKQNKYISIEHRRKKIKIAGLKMDHNIITATGVKINIALSTIIL  
SKYKISHIINGVAGSIYSDKNKPIKIGDVIVSTETSYDFDLRFEYELGVHEHR  
KEKATVALIRKTSKIKINNIITSYMGLIITGDFIDHGFPEIPEEPENAIADIMESA  
MAVAVGKIFPIIFIIIRGISDIYNNENNVDYKFKLAKSSSSAKIVENLIKLM"  
complement(3343. .3441)  
/gene="BB107"  
complement(3343. .3441)  
/gene="BB107"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/product="B. burgdorferi predicted coding region BB107"  
/protein\_id="AAC66215.1"  
/db\_xref="GI:2690112"  
/translation="VQLKGAIIKLEDYERIKRIQLNLIKLIH"  
3576. .3755  
/gene="BB108"  
3576. .3755  
/gene="BB108"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="B. burgdorferi predicted coding region BB108"  
/protein\_id="AAC66196.1"  
/db\_xref="GI:2690093"  
/translation="MISFPDYNPSASNIFKITTPKCLAKKLIPLIQRKQFNLES  
FKNKIKYISLWIGK"  
3745. .3882  
/gene="BB109"  
3745. .3882  
/gene="BB109"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="B. burgdorferi predicted coding region BB109"  
/protein\_id="AAC66197.1"  
/db\_xref="GI:2690094"  
/translation="MASEKKIDAEKSNEKVSQKKSNKNKNIYEKDTLGLVKIKKSS  
I"  
3911. .4315  
/gene="BB110"  
3911. .4315  
/gene="BB110"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="B. burgdorferi predicted coding region BB110"  
/protein\_id="AAC66198.1"  
/db\_xref="GI:2690095"  
/translation="MGRAMTPHQOALKVNLNGLKMIINLDELKITAIDGDDNNISKKK  
NOLKRVYDNNQNSLIKYLTKNSSNEKIEKKCIKTLMHVVEYIFESVODELKKKK  
KHEHNSILITLWEAIKIKNSAIAIHVCENID"  
complement(4623. .4721)  
/gene="BB111"  
complement(4623. .4721)  
/gene="BB111"  
/note="hypothetical protein; identified by Glimmer;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="B. burgdorferi predicted coding region BB111"  
/protein\_id="AAC66199.1"  
/db\_xref="GI:2690096"  
/translation="MSKLOGYSTVNVVFPKIQIKILIVNKGSVLF"  
5128. .5346  
/gene="BB112"

CDS		5128..5346	/gene="Bil12"	/note="hypothetical protein; identified by Glimmer;
			putative	
			/codon_start=1	
			/transl_table=11	
			/product="B. burgdorferi predicted coding region Bil12"	
			/protein_id="AC66200.1"	
			/db_xref="GI:2690097"	
			/translation="MIKFPKNHSHIKIIEYEDVTIKMDREVSLEFKLHKNTLED WLEYOKEEMOKIKEFANKFIKKRPKI"	
gene		5609..5707	/gene="Bil13"	
		5609..5707	/gene="Bil13"	
CDS		5609..5707	/gene="Bil13"	
Query Match	11.0%	Score 82.2;	DB 1;	Length 27323;
Best Local Similarity	49.9%	Pred. No.	0.00013;	
Matches 333;	Conservative	0;	Mismatches 308;	Indels 26; Gaps 4;
Db	13	CAATTAAGATATTGTTAAAGAATTGAATTGCTAATTATTTATGGTAAATCAGAGACT	72	
Db	20290	CTAATAGAGGATTCCTATGTATCATCAAAAAGATATATTAAATCAAAATAACCTAT	20231	
Qy	73	CTATTGGTAACGCAATTCGAGTACAACTTTGAAAAAGTTAAATTTAAATGT-TCOA	131	
Db	20230	TTATTATACAACCTCATCTTAAACCCTTAATGAAATTTATAGTTAAGTATTCCTTGG	20171	
Qy	132	AAGCCTTTTTTAAATTTATATTATGCTACCA-TAATACCGATTAAATAAGGGGTTT	190	
Db	20170	GCGTTTGTGATTAANTTGATTAATTTAATTAATTAATAATTTATGGAGGTAATATCA	20111	
Qy	191	TATGATTAAGTTTTTAATTTTATTTCGGTAATCTTTTGTGCTTTTCTAGTTTGCICA	250	
Db	20110	TATGAAAAAATTTTAACTATGATTTTATTTTGGCTTAAACAATAAACAAATCTTGGCAC	20051	
Qy	251	AGATATTTCTAAACACACTTTTAATCTGGAGCGGAGAAATTTTGGTTATGAAC	310 /	
Db	20050	ACAAAG---TANGCTGAAAAAAGTGTGGAGATTTGAACCATTTAGAAATATAAAG	19994	
Qy	311	TAAATGAAGAAGTCTCTGTGACATTTTATTTGAACCTTTTTAGGGTTCGGATAG	370	
Db	19993	CGAAAAAGCACTATATCTAGCCGCCATTCCTTTGAAATTAATTTAACTTTAAGGAATAG	19934	
Qy	371	TTCTTTGCTCAAGGAGATATCCCTGGAGGCTCTCTTATCTTGGATTGATGCGGTTG	430	
Db	19933	ATCTTTGTCCAAGAGATTAATTGCGTGTGCGCAGCTTGGATCTCACTTAATTAG	19874	
Qy	431	TATAGGTTAATACTTACAGAGCTTAATTTGACATCAAGAVTTTGTATTAATGCTAA	490	
Db	19873	AGGAATACTTTGCAATGCTGAAATATCTTGCCCATACAGATGATGAAMCAAGGCAAC	19814	
Qy	491	AAAACCTGATTTTAAGTGACTTGGGGTTAAGGAATGATCTTGGCAGGCTGATTAAT	550	
Db	19813	AACTGGGCAATAT-----AATACACGATNAGAGTATGCGCAGAT	19775	
Qy	551	GGCTGTGACAAGATTGACAGAAATTTGCTTCATTTACATTTGCTAATTAATTAACAG	610	
Db	19774	TATACATCCCCACATAGCCTCACTTAATTTTCATTTACATTTGCAAAATAACACAAATC	19715	
Qy	611	GAACTGAAAAATAGCTTAATATATAGCCTTGGGAGAGATTGAGCCTAGTTTGATATTAA	670	
Db	19714	AAATCTTAAAAAAGACATCGCATGTGATATTCGCGGTTTGAACCAATTTGATATTGG	19655	
Qy	671	CATGGG 677		
Db	19654	AATTAAG 19648		
RESULT 8				
LOCUS	AC117070	156533 bp	DNA	linear HTG 06-APR-2001
DEFINITION	Dicyostelium discoideum chromosome 2 map 2268371-2355204 strain			

```

ACCESSION      AC117070.1
VERSION        GI:20066187
KEYWORDS       HTG; HTGS PHASE2.
SOURCE         Dictyostelium discoideum.
ORGANISM       Dictyostelium discoideum.
REFERENCE      Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
AUTHORS        1 (bases 1 to 156533)
               Gloeckner,G., Eichinger,L., Szafarski,K., Pachebat,J., Dear,P.,
               Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
               Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
               Tungal,A.A.
TITLE          Sequence and Analysis of Chromosome 2 of Dictyostelium
JOURNAL        Unpublished
REMARK         The Dictyostelium Genome Sequencing Consortium
REFERENCE      2 (bases 1 to 156533)
AUTHORS        Baumgart,C.
TITLE          Direct Submission
JOURNAL        Submitted (06-APR-2002) Genome Analysis, Institute of Molecular
               Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT        CDS predictions from GeneID may contain errors. Further information
               is available from IMB Jena, Department of Genome Analysis
               (http://genome.imb-jena.de/dictyostelium/)
               and the University Cologne, Institute for Biochemistry I
               (http://www.uni-koeln.de/dictyostelium/project.shtml)
               Funding
               Agency : Deutsche Forschungsgemeinschaft (DFG).
               * NOTE: This is a 'working draft' sequence.
               * This sequence will be replaced
               * by the finished sequence as soon as it is available and
               * the accession number will be preserved.
FEATURES       source
               location/Qualifiers
               1..156533
               /organism="Dictyostelium discoideum"
               /strain="AX4"
               /db_xref="taxon:44689"
               /chromosome="2"
               /map="2268371-2355204"
               join(3827..3945,5512..5557)
               /note="ORF_ID:dd_00443"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="AA009328.1"
               /db_xref="GI:20066214"
               /translation="MSADQPPRLTMLMSADQPPQIQISMSVDEPTLITHSVGLG
               ASLDGNGOTA
               complement(7554..8315)
               /note="ORF_ID:dd_00439"
               /codon_start=1
               /product="Discoidein I, D chain"
               /protein_id="AA009325.1"
               /db_xref="GI:20066211"
               /translation="MSTQGLVTLGNAOCHLRTSTNGVHTQFNALNYKNGKNTI
               DGEAMCSSIYDINOYIVAGCEVPRTFMCVALOGGDHDOVTSYKIRYSLDNVWSE
               YRNGAIIYGVDRNTVNVHFFDTPIRASIAIHPLTNNHISLRCEFTOPQSSVTO
               VGADIIYDNCALNMGSKREVVPYKQFEPAFLPKVALNFDQIDCDTDAINQTRIGV
               QPRNITTKGFPCVFTVMNKVYSLRADYIAFALE"
               complement(9039..9800)
               /note="ORF_ID:dd_00438"
               /codon_start=1
               /product="Discoidein I, C chain and B chain."
               /protein_id="AA009326.1"
               /db_xref="GI:20066212"
               /translation="MSTQGLVTLGNAOCHLRTSTNGVHTQFNALNYKNGKNTI
               DGEAMCSSIYDINOYIVAGCEVPRTFMCVALOGGDHDOVTSYKIRYSLDNVWSE
               YRNGAIIYGVDRNTVNVHFFDTPIRASIAIHPLTNNHISLRCEFTOPQSSVTO
               VGADIIYDNCALNMGSKREVVPYKQFEPAFLPKVALNFDQIDCDTDAINQTRIGV
               QPRNITTKGFPCVFTVMNKVYSLRADYIAFALE"
               join(12482..12692,13544..13672,14278..14315)
               /note="ORF_ID:dd_00437"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="AA009327.1"
               /db_xref="GI:20066213"
               /translation="MIKEEDPOLIKREKYSISIEDLDYEDSDCLDCGQPKYHSRHC
               SSIHOAVNNKVMCHCKRYPRESPRADAEDRRKRIYKQRLQRELQDTRRESINR
               TASIKRLREKRYVASIPILKN"
               /note="ORF_ID:dd_00448"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="AA009329.1"
               /db_xref="GI:20066215"
               /translation="MNFEDKLGEINLCEELNKKPTQIRLEPYPTLELNSSTRLID
               NFIVGTLNIVRQKGIYVFGRIILKEKFTLHCECYSTFEDEPQOKFLYIPGVHYV
               HPTVFNKIKKINNOVXENGLPELIDECNIYMLPDVYISGLIFENKEDDSII
               KNIIDAKGSKHPIRAPITSEYNNKEIRQVKKNNKNOQPTTNTTRKPIPSPTPT
               NASTPTPTTPTTPTTPTTPTTPTTITTHQLPSTNTIINLNKTTTTTTTTTTTKPK
               ITTSNGTFPIASISIRK"
               /note="ORF_ID:dd_00447"
               /codon_start=1
               /product="Putative Retron-type reverse transcriptase"
               /protein_id="AA009330.1"
               /db_xref="GI:20066216"
               /translation="MNFPERIDHSGSGGTGAIENDRTRKHISINFRDDGRILSIK
               YNFSNINILINAPATISERTPTIINSKLSLKKYNSINHOIAGDPNNHDCNSFEG
               TELRKIIDQMDLIDGIEHNPPTPPRSKRRIDRIYCHPLLQNSKLVHNVFNKSD
               HPITITITOTNRETTTTTTTTTKLERLPWLCKEILANKHIHGLSELISKNKDKTSV
               EMTKFNKNVIRDYLRKEDONKIKERKNRKYVTHKLDSGDIIPMRKELNEISIRIL
               BEERVKAMVDIKLHLHOETPSKYLTSLKSRADKSLFQDKDNNHDSKNSIAK
               REVEFYODQYEKEEDNEETHKLLKEWEVDVLIKLEIDRIRIHEVYKAKTSSII
               KSPGLDGINALEYKYHINSIARILITFAFDLTNKEKEIPTRKEGVITFEKGEELN
               ISNRPTLLNMDYKILSLIINSRLDITSKINRFGNGVFNRIQDNIOTMKVIE
               ISNRKNNTLITFYPNKAFDISIKSISTRTLEHGIGTPKFTAILILNKDTRKIKI
               NDELVNGITTRTGOGDTPPTIFALVLEPLIDINDNTKGTLLPNPSKSLKLTAF
               ADDIATFNSTSEELMKINTKIOKYSATSSSNKERTYVIAIGDPHDLPEQESTVPE
               RYLGLNFTKGLNSKRYNTLIQEMKNMLIKSQAATMKAKTILKTYVLSKLTYSQYM
               DNLNEQIIEIENNTRMWFLFSSVKNYVEERKYTKMMKIDRAYADMKGGIKLMDIEL
               RHIAKIWMNRLNHNNTNNNNNTLOEYEMEDLSKKKPTLQDSTICADPTGYSESNTF
               KTQR"
               join(18321..18339,18404..18924)
               /note="ORF_ID:dd_00446"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="AA009331.1"
               /db_xref="GI:20066217"
               /translation="MAEYWGGRNTLPREFRSRLPQINHERDTRCKICGHLFDPYSHL
               FTLCODIIDIEKTIISTVNLKSFIRKIQMSMDTLDISKYNRTERJEPNLSIGIIAOLW
               KIICHKLFNTDESKPEPKFEQVIEITELLNLIETEKFTLKIKHDEAILKNTQDLH
               KYXENKAMOTPAAPMLPI"
               /note="ORF_ID:dd_00444"
               /codon_start=1
               /product="Putative Nuclear protein"
               /protein_id="AA009332.1"
               /db_xref="GI:20066218"
               /translation="MNNIFLISSIIISIVFVSGEALIDSSSEISVGLISDSQACLA
               SNQKRVIIQWSSGGQIVNSNTSSVSAEQAQFNDVIDIYALFCSKCDENYPRSSAIQS
               LVSSKLDGQINENMLMIDVQDCDGGCAGSADADVADYVQEAQGLGVLYVYSSEGE
               MPQYGNLSLISQYPLNAYHYDNDNPSFDITAFEEYGGTSPSMKQYIANTNOCQGVSD
               LDYVSGSGCSTSSGASGASGASGASGASGSSNSGSSNSGSSNSGSSNSGSSNS
               GSGNSGSSNSGASGSGSGSSI"
               complement(join(21433..23231,23387..23705))
               /note="ORF_ID:dd_00442"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="AA009333.1"
               /db_xref="GI:20066219"
               /translation="MSINKKOKITINEINERKNDISSGNNIIRKTEGNEKCTD
               IPTPIPTIIDELISLPSPEPPPTTIDNNKTPITITPQIDLDPSATLILSP
               INSTIILGSHVHIRGSSDESEILIRKKPDTVEVLELCSRGAILFNNNDSSDYFN
               NNNNNNNNNNTSSKSRTRKILIDGRTKSVYSRFLRQKHQHLISLDLDQDSASTSK
               FVLETEIODIDIEDMTQEDNELYTSEETQEDENFDDDDYGVGFYLDHNDNINILNM
               OPRNFNPLYDDNDLIEBNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
               NKSNDNLKTKTLINNNKNNIIFNSDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

```







LOCUS	AX251990	AX251990	Sequence	251 from Patent WO0168911.	7110 bp	DNA	linear	PAT 05-OCT-2001
DEFINITION								
ACCESSION	AX251990							
VERSION	AX251990.1	GI:15985345						
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
FEATURES								
source								
BASE COUNT								
ORIGIN								
Query Match								
Best Local Similarity								
Matches								
35	AATTGAAATGCTAATTTTATGCTTAAATCAAGAAGCTCTATTGCTTAAGCGAATTTGAG	8.8%; Score 66; DB 6; Length 7110;						
1672	AATTGAAATTTAAAAAGATATATTTTTGAATTTTTTTAGTTTAAAGTAGAGTTATTA	51.0%; Pred. No. 0.08;						
95	TACAAATTTGAAAAAGTTAAATTTAAATAGTCCAAAGCCTTTTAAATTCATTAA	0; Mismatches 150; Indels 0; Gaps 0;						
1732	TTTGGGTGAAAGCAAAAAGAAATTAATAATTGSAAGAGATGATTTAGTATTTT							
155	TATGCTACCAATATACAGTTTAAATTAAGGGTTTATTAAGATAAGTTTAAATTTTAT							
1792	AGAAATTAATTAATTAATTTAGGAATACAGAGAGATTTTATTTAATTTTAAATTAAT							
215	TTTGGTAACTCTTTTGTGCTTTTCTAGCTTTTGCTCAACATGATTCCTAAAGCATT							
1852	TTGATGTGTTGTTGGGTGTTAAATTTGTTAAATTTGATGATTTTAAATTTTAA							
275	TCTGGAGCGGAGAAAAATTTTGGTTTATGAACATATATAAGAAATTCCTGTAC							
1912	TTAGAGAAAGTAGAGGATTTGATGTTTAGAGAAAGTAGTAAATTTTATTTTATTTT							
335	ATTATT 340							
1972	TTTTTT 1977							
RESULT 11								
LOCUS	AX346458							
DEFINITION								
ACCESSION	AX346458							
VERSION	AX346458.1	GI:18494344						
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
FEATURES								
source								
ORIGIN								
Query Match								
Best Local Similarity								
Matches								
35	AATTGAAATGCTAATTTTATGCTTAAATCAAGAAGCTCTATTGCTTAAGCGAATTTGAG	8.8%; Score 66; DB 6; Length 7110;						
1672	AATTGAAATTTAAAAAGATATATTTTTGAATTTTTTTAGTTTAAAGTAGAGTTATTA	51.0%; Pred. No. 0.08;						
95	TACAAATTTGAAAAAGTTAAATTTAAATAGTCCAAAGCCTTTTAAATTCATTAA	0; Mismatches 150; Indels 0; Gaps 0;						
1732	TTTGGGTGAAAGCAAAAAGAAATTAATAATTGSAAGAGATGATTTAGTATTTT							
155	TATGCTACCAATATACAGTTTAAATTAAGGGTTTATTAAGATAAGTTTAAATTTTAT							
1792	AGAAATTAATTAATTAATTTAGGAATACAGAGAGATTTTATTTAATTTTAAATTAAT							
215	TTTGGTAACTCTTTTGTGCTTTTCTAGCTTTTGCTCAACATGATTCCTAAAGCATT							
1852	TTGATGTGTTGTTGGGTGTTAAATTTGTTAAATTTGATGATTTTAAATTTTAA							
275	TCTGGAGCGGAGAAAAATTTTGGTTTATGAACATATATAAGAAATTCCTGTAC							
1912	TTAGAGAAAGTAGAGGATTTGATGTTTAGAGAAAGTAGTAAATTTTATTTTATTTT							
335	ATTATT 340							
1972	TTTTTT 1977							
RESULT 11								
LOCUS	AX346458							
DEFINITION								
ACCESSION	AX346458							
VERSION	AX346458.1	GI:18494344						
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								

BASE COUNT	1749 a	289 c	1822 g	3250 t
ORIGIN				
Query Match	8.8%;	Score 66;	DB 6;	Length 7110;
Best Local Similarity	51.0%;	Pred. No. 0.08;		
Matches 156; Conservative	0;	Mismatches 150;	Indels 0;	Gaps 0;
QY	35 AATTGAATTCGTAATTTTATGCTTAATCAAGAAAGCTCATTTGGTAAGCGAATTTCCAG	94		
Db	1672 AATTGAATTTTAAAAAGATATTTATTTTGAAAATTTTTTTTAGCTTAGAAGGCTAGTTA	1731		
QY	95 TAACAATTTGAAAAAGTTAAATTTTAAATVAGTCCAAAACCCCTTTTAAATTCATTAA	154		
Db	1732 TTTGGGTGAAAGGAAAAAGAAAATAATTAATTTGAAAAGGATGTAATTAATTTT	1791		
QY	155 TATGCTACCATAATACCAAGTTAAATAAAGGGCTTTTATGATTAAGTTTAAATTTTAT	214		
Db	1792 AGAATTAATTTAAATTTTAAAGCAATACAGAGAGATTTTATTTTAAATTTTAAATTAAT	1851		
QY	215 TTTCGTAATCTTTTGTCTTTTCTTAGATTTTGCTCAAGATGATTCATAAGCATTTTAA	274		
Db	1852 TTGATGTTGTTTGTGGGTGTTAAATATTTGTTAAATGATGATTTTTTATTTATTA	1911		
QY	275 TCTGGAGCGGAGAAAAATTTTGGTTATGAACCTATAGAAGATTCCTGTGCC	334		
Db	1912 TTAGAGAGAGATGAGGCGATTTGATGTTTAGAAGATGTAATAATTTATTTTATTTT	1971		
QY	335 ATTTT 340			
Db	1972 TTTT 1977			
RESULT 12				
LOCUS	AX349019	7110 bp	DNA	Linear PAT 06-FEB-2002
DEFINITION	Sequence 477 from Patent WO0202807.			
ACCESSION	AX349019			
VERSION	AX349019.1 GI:18615054			
KEYWORDS	.			
SOURCE	synthetic construct.			
ORGANISM	synthetic construct			
REFERENCE	artificial sequences.			
AUTHORS	Olek,A., Piepenbrock,C. and Berlin,K.			
TITLE	Diagnosis of diseases associated with cell signalling			
JOURNAL	Patent: WO 0202807-A 477 10-JAN-2002;			
FEATURES	Epidemiomics AG (DE)			
source	Location/Qualifiers			
	1..7110			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="chemically treated genomic DNA (Homo sapiens)"			
BASE COUNT	1749 a	289 c	1822 g	3250 t
ORIGIN				
Query Match	8.8%;	Score 66;	DB 6;	Length 7110;
Best Local Similarity	51.0%;	Pred. No. 0.08;		
Matches 156; Conservative	0;	Mismatches 150;	Indels 0;	Gaps 0;
QY	35 AATTGAATTCGTAATTTTATGCTTAATCAAGAAAGCTCATTTGGTAAGCGAATTTCCAG	94		
Db	1672 AATTGAATTTTAAAAAGATATTTATTTTGAAAATTTTTTTTAGCTTAGAAGGCTAGTTA	1731		
QY	95 TAACAATTTGAAAAAGTTAAATTTTAAATVAGTCCAAAACCCCTTTTAAATTCATTAA	154		
Db	1732 TTTGGGTGAAAGGAAAAAGAAAATAATTAATTTGAAAAGGATGTAATTAATTTT	1791		
QY	155 TATGCTACCATAATACCAAGTTAAATAAAGGGCTTTTATGATTAAGTTTAAATTTTAT	214		
Db	1792 AGAATTAATTTAAATTTTAAAGCAATACAGAGAGATTTTATTTTAAATTTTAAATTAAT	1851		
QY	215 TTTCGTAATCTTTTGTCTTTTCTTAGATTTTGCTCAAGATGATTCATAAGCATTTTAA	274		
Db	1852 TTGATGTTGTTTGTGGGTGTTAAATATTTGTTAAATGATGATTTTTTATTTATTA	1911		

Db 1852 TTGATGCTGTGTTGGGTTGCTTAATATTTGTTAATATGATGATTTTATTTATTA 1911  
QY 275 TCTGGAGCGGAGAAAAATTTGTTTATGAACATAAAGAAATCTCTGAC 334  
Db 1912 TTAGAGAAAGTAAAGGATTTGATGTTTGAAGAGTAAATTTATTTTATTTT 1971  
QY 335 ATTTT 340  
Db 1972 TTTT 1977

RESULT 13  
AX458543  
LOCUS AX458543 34688 bp DNA linear PAT 08-JUL-2002  
DEFINITION Sequence 89 from Patent WO024654.  
ACCESSION AX458543  
VERSION AX458543.1 GI:21725207  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1  
AUTHORS Schacht, O.  
TITLE Diagnosis of diseases associated with angiogenesis  
JOURNAL Patent: WO 024654-A 89 13-JUN-2002;  
Epidemiol. Infect. 126: 1-11 (2001)  
FEATURES  
source 1. 34688  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 9920 a 444 c 6941 g 17383 t  
ORIGIN

Query Match 8.8%; Score 66; DB 6; Length 34688;  
Best Local Similarity 51.0%; Pred. No. 0.057;  
Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 35 AATGAAATGCTAATTTATGTTAAATCAAGAGCTCTATGTTAGCAATTCGAG 94  
Db 1672 AATGAAATTTTAAAGTAAATTTTGAATTTTATGTTAGTATGTTAGTAT 1731  
QY 95 TAACAATTTGAAAAAGTAAATTTAATAGTTCCAAAAGCCTTTTAAATTCATTA 154  
Db 1732 TTTGGGTAAGAAAGAAAGAAATTAATGAAAGAGATGATTTATTTT 1791  
QY 155 TATGCTACCAATTAACAGGTTTAAATGAAGGTTTATGAATAAGTTTAAATTTAT 214  
Db 1792 AGAATATATTTAATTTAGTAAGTAAGAGAGATTTTATTTATTTTAAATTTAT 1851  
QY 215 TTTGGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATTTCAAGCATT 274  
Db 1852 TTGATGCTGTGTTGGGTTGTTAAATTTGTTAATTTGTTATTTTATTTATTA 1911  
QY 275 TCTGGAGCGGAGAAAAATTTGTTTATGAACATAAAGAAATCTCTGAC 334  
Db 1912 TTAGAGAAAGTAAAGGATTTGATGTTTGAAGAGTAAATTTATTTTATTTT 1971  
QY 335 ATTTT 340  
Db 1972 TTTT 1977

RESULT 14  
AX348445  
LOCUS AX348445 7441 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 140 from Patent WO0202806.  
ACCESSION AX348445  
VERSION AX348445.1 GI:18614481  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source 1. 7479  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Method and nucleic acids for pharmacogenomic methylation analysis  
JOURNAL Patent: WO 0202806-A 140 10-JAN-2002;  
Epidemiol. Infect. 126: 1-11 (2001)  
FEATURES  
source 1. 7441  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 2128 a 66 c 1267 g 3980 t  
ORIGIN

Query Match 8.5%; Score 63.8; DB 6; Length 7441;  
Best Local Similarity 46.9%; Pred. No. 0.18; Mismatches 262; Indels 3; Gaps 1;  
Matches 234; Conservative 0; Mismatches 262; Indels 3; Gaps 1;

QY 17 TAAATATGTTTAAAGATTTGAATTTGCTAATTTTATGTTTAAATCAAGAGCTAT 76  
Db 6369 TGAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6428  
QY 77 TGTAAAGCAATTTGAGTAAATTTGAAAAAGTTAAATTTAAATTTTAAATTTTAAATTT 136  
Db 6429 TAGTAAAGCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 6488  
QY 137 TTTTAAATTTTCAATTA--ATATGCTAACATAATTTTAAATTTTAAATTTTAAATTT 193  
Db 6489 TTTAATATTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 6548  
QY 194 GAATTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 253  
Db 6549 TTTTCGATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 6608  
QY 254 TGAATTTTAAAGCAATTTTAAATTTGAGGAGCGGAGAAAAATTTTGGTTTAAATTTAA 313  
Db 6609 TATGCTTTGGAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 6668  
QY 314 TAAGAAATTTCTCTGTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 373  
Db 6669 TTTTAAAGATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 6728  
QY 374 TTTTCTCAAGAGATTTCTTGGAGTTTCTGATTTTATTTTATTTTATTTTATTTTATTT 433  
Db 6729 TATAGCTTTAAAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6788  
QY 434 AGGTTTAAATTTTAAAGCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 493  
Db 6789 TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 6848  
QY 494 AGCTGATTTTAAAGTGAAT 512  
Db 6849 TTAGATTTTAAAGTGAAT 6867

RESULT 15  
AX339173  
LOCUS AX339173 7479 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 40 from Patent WO0176451.  
ACCESSION AX339173  
VERSION AX339173.1 GI:18129250  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with metabolism  
JOURNAL Patent: WO 0176451-A 40 18-OCT-2001;  
Epidemiol. Infect. 126: 1-11 (2001)  
FEATURES  
source 1. 7479  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

```

BASE COUNT      2133 a      68 c      1273 g      4005 t
ORIGIN

```

Query Match	8.5%	Score	63.8	DB	6	Length	7479
Best Local Similarity	46.9%	Pred. No.	0.18				
Matches	234	Conservative	0	Mismatches	262	Indels	3
						Gaps	1

OY	17	TAAAGATTTGGTAAAAAGAAATTCMAATTCGTAATTTTATGGTTAATCAAGACGTCTAT	76
Db	6369	TGAAATATTTTTTTTTTACGTATTTTCTTTTACGTTTGTGTTTTAGCTTAGCTTCGTTT	6428
OY	77	TGCTAACCCGAATTTCCAGTACACATTTGAAAAAGTTAAATTTAAATAGTCCAAAGCC	136
Db	6429	TAGGTAGAGATATTTTTTAAAAATTTTATATATTTTTTATATTTTATTTTATAGTA	6488
OY	137	TTTTTTTAAATTCATTA--ATATGTCACATATATCCAGTTATATAAGGGGTTCAT	193
Db	6489	TTTTAATATATATTTATAGATTTTTTTTAAAGGTTTTTAATNAGGTTTTTATTTTTTTTAT	6548
OY	194	GAAATACGTTTATTAATTTATTTTGGTAATCTTTGTGCTTTTCTACTTTTGCTCAAGA	253
D	6549	TTTTTCGATTTTTTATTTTGAATAATGTAAATTTTTTTTTTTTTCGTATTTGATTAATTT	6608
OY	254	TGATTCAAAAGACGCTTTAATCTCGGAGCGGAGAAAAATTTTGGTTTATGAAACTAA	313
Db	6609	TATTCGTTGGGATATTTTATATTTTCCGTTATAGAGTATTTATATTTTATTTATATAT	6668
OY	314	TAAAGAAAGATTCCTGTGTACATTTTTTATGAACTTTTATTTAGGGTTCGGGATAGGTC	373
Db	6669	TTTTTAAAGACTTATTTCTGTTTTTATAGAAAGTATGATTTTATTTTAAAGGTGTTT	6728
OY	374	TTTTGTCACAGAGATATTCCTTGAGGTCCTATATCTTGATTTGATGGCGTGGGAT	433
Db	6729	TATAGGTTTAAAGTTATTTATTTTTTTTTTTTTTTTATTAAGTATATTTATTTTGTTTAATAA	6788
OY	434	AGGGTTATATCTACTACAGAGACTTTTGGACATCAAGATTTTGCATTAATATGCTAAAAA	493
Db	6789	TTTTTTTTTAAATATATATTAAGTTTATTTGTTAAAAAGTTAGATGCAATTTATTTATTTAGT	6848
OY	494	AGCTGATTTTAACTGGAGCT	512
Db	6849	TTAGGATTTTAAAGTTATTT	6867

Search completed: March 27, 2003, 03:12:31  
Job time : 1747.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:05:42 ; Search time 165.602 Seconds  
(without alignments)  
10185.565 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749  
Sequence: 1 tagaatttcacaataaa.....attatctagaataatgggtg 749

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	749	100.0	749	20	AAX30100
2	607	81.0	862	20	AAX30099
3	558.2	74.5	759	20	AAX30098
4	402.8	53.8	540	20	AAX61505
5	368.8	49.2	477	20	AAX61506
6	98.4	13.1	1761	20	AAX20310
7	91.6	12.2	9399	20	AAX20261
8	82.2	11.0	7074	20	AAX20266
9	66	8.8	7110	22	AAS46489

10	66	8.8	7110	24	ABL70587	Chemically treated
11	66	8.8	7110	24	ABL33556	Human immune syste
12	66	8.8	7110	24	AAS61290	Human immune syste
13	66	8.8	34688	24	ABO67059	Human angioegenis
14	63.8	8.5	7441	24	ABK40058	Human chemically p
15	63.8	8.5	7479	24	AAS63345	Human chemically pretrea
16	63.8	8.5	14095	24	ABL32476	Human immune syste
17	63.2	8.4	6131	24	ABL32891	Human immune syste
18	62.2	8.3	6391	24	ABK28137	DNA transcription
19	61.8	8.3	14023	24	ABL34105	Human immune syste
20	60.2	8.0	9707	24	ABL33421	Human immune syste
21	59	7.9	3873	24	ABO67089	Human angioegenis
22	59	7.9	5675	24	ABL33276	Human immune syste
23	59	7.9	6464	24	ABL33515	Human immune syste
24	58.6	7.8	9731	24	ABL32990	Human immune syste
25	58.4	7.8	8237	22	AAS46802	Tumour suppressor
26	58	7.7	15387	24	ABL32184	Human immune syste
27	57.6	7.7	7781	24	ABL33169	Human immune syste
28	57.2	7.6	13125	24	ABL70283	Chemically treated
29	57.2	7.6	13125	24	ABL33226	Human immune syste
30	57.2	7.6	13125	24	ABL34556	Human immune syste
31	57	7.6	4163	24	ABL34330	Human metastasis a
32	57	7.6	7348	22	AAS46336	Human immune syste
33	56.8	7.6	6244	24	ABL32484	Tumour suppressor
34	56.8	7.6	8077	24	ABL33176	Human immune syste
35	56.6	7.6	9021	24	ABK33988	Human immune syste
36	56.6	7.6	9021	24	ABK40075	Human DNA for stag
37	56.6	7.6	9021	24	ABL34232	Human immune syste
38	56.4	7.5	19380	22	AAS61426	Human immune syste
39	56	7.5	6022	22	AAS46661	Tumour suppressor
40	56	7.5	17594	24	ABL34026	Human immune syste
41	55.4	7.4	14147	22	AAS46743	Tumour suppressor
42	55.4	7.4	14147	24	ABK33955	Human DNA for stag
43	55.2	7.4	6169	22	AAS46370	Tumour suppressor
44	55.2	7.4	6169	24	ABN80097	Human chemically m
45	55.2	7.4	11538	24	ABL33156	Human immune syste

#### ALIGNMENTS

RESULT 1	
AAX30100	standard, DNA; 749 BP.
ID	AAX30100
XX	
AC	AAX30100;
XX	
DT	17-JUN-1999 (first entry)
XX	
DE	Borrelia garinii IP90 protein encoding cDNA.
XX	
KW	Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.
XX	
OS	Borrelia garinii.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	192..725
XX	/*tag= a
XX	
PN	W09912960-A2.
XX	
PD	18-MAR-1999.
XX	
PF	04-SEP-1998; 98WO-1B01424.
XX	
PR	16-SEP-1997; 97US-0059036.
PR	10-SEP-1997; 97DK-0001041.
XX	
PA	(SYMB-) SYMBICOM AB.
XX	
PI	Bergstroem S;
XX	
DR	WPI; 1999-215027/18.

DR P-PSDB; AAY04280

PT	Nucleic acid	from <i>Borrelia burgdorferi</i>	encoding virulence associated
PT	protein	P13	

PS Claim 1; Page 113-114; 118pp; English.

CC The present sequence encodes a Borrelia garinii IP90 protein. The  
CC present invention describes an isolated nucleic acid (I) that: (i)  
CC encodes a polypeptide fragment (II) immunologically reactive with  
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of  
CC Borrelia burgdorferi B313, but not reactive with most proteins from  
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions  
CC to specific nucleic acid sequences but not to genomic DNA from most  
CC other spirochaetes. (I), (II) and transformed cells are useful in  
CC vaccines to protect against infection by B. burgdorferi sensu lato.  
CC (I), (II) and antibodies raised against (II) are used to diagnose such  
CC infections. In standard immunoassays or amplification/hybridization  
CC tests, (I) are also used to produce recombinant (II). The 13 kD protein  
CC is involved in virulence and is highly conserved within B. burgdorferi  
CC sensu lato, but is absent from other Borrelia species (e.g. those  
CC responsible for relapsing fever or yaws borreliosis). It should provide  
CC a more specific and sensitive antibody response, and diagnosis,  
CC compared with use of whole bacteria as antigen.

Sequence 749 BP; 234 A; 82 C; 150 G; 283 T; 0 other;

Query Match	100.08;	Score 749;	DB 20;	Length 749;
-------------	---------	------------	--------	-------------

Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TAGAAATTTTCAACAAATTAATATTTGTGTAAACAAATTGAATTTGCTAAATTTATGGTTA	60
Db	1	TAGAAATTTTCAACAAATTAATATTTGTGTAAACAAATTGAATTTGCTAAATTTATGGTTA	60
QY	61	AATACAGAGACTCTATTGGTGTACGGCAATTTGAGTAACAATTTGAAAAAGTTAAATTTTA	120
Db	61	AATACAGAGACTCTATTGGTGTACGGCAATTTGAGTAACAATTTGAAAAAGTTAAATTTTA	120
QY	121	AATAGTTCCAAACCCCTTTTCTTAATTTTCATTAATATATGCTACCCCTTAATACAGTTTATA	180
Db	121	AATAGTTCCAAACCCCTTTTCTTAATTTTCATTAATATATGCTACCCCTTAATACAGTTTATA	180
QY	181	AAGGGGTTTTTATGAAATPAGTTTTTAAATTTTAAATTTTGGTAATCTTTTGTCTTTTCTTA	240
Db	181	AAGGGGTTTTTATGAAATPAGTTTTTAAATTTTAAATTTTGGTAATCTTTTGTCTTTTCTTA	240
QY	241	GTTTTTGCACAGATGATTTCTTAAGCACTTTTAATCTGGAGGGGAGAAAAATTTTGG	300
Db	241	GTTTTTGCACAGATGATTTCTTAAGCACTTTTAATCTGGAGGGGAGAAAAATTTTGG	300
QY	301	TTTATGAAACTAATPAGAAAGATTTCTGTGACCAATTTTATGAAACCTTTTATAGGGT	360
Db	301	TTTATGAAACTAATPAGAAAGATTTCTGTGACCAATTTTATGAAACCTTTTATAGGGT	360
QY	361	TCCGGATPAGGTTCTTTTGGCTTAAGGAAATATCCTTGGAGGTTCTTATTTCTTGGATTG	420
Db	361	TCCGGATPAGGTTCTTTTGGCTTAAGGAAATATCCTTGGAGGTTCTTATTTCTTGGATTG	420
QY	421	ATGCGGTTGGTATAGGGTTAATACCTTAACAGAGCTTATTTGSAACATCAAGAAATTTTGATA	480
Db	421	ATGCGGTTGGTATAGGGTTAATACCTTAACAGAGCTTATTTGSAACATCAAGAAATTTTGATA	480
QY	481	ATAATGCTCAAAAAAGCGTATTTTAAATGAGGACTTGGGGTAAAGGAATATGTTGGCAGGTG	540
Db	481	ATAATGCTCAAAAAAGCGTATTTTAAATGAGGACTTGGGGTAAAGGAATATGTTGGCAGGTG	540
QY	541	TGGTTACTATGCGTGTGCAAGATTGACAGAAATTTGTTCTTCCATTTACATTTGGCTATA	600
Db	541	TGGTTACTATGCGTGTGCAAGATTGACAGAAATTTGTTCTTCCATTTACATTTGGCTATA	600
QY	601	ATTATACAGGAGCTGAAAAATATGCTTAAATATATAGCCTTGGGAGATTTTGAGCCTAAGTT	660
Db	601	ATTATACAGGAGCTGAAAAATATGCTTAAATATATAGCCTTGGGAGATTTTGAGCCTAAGTT	660

Db 601 ATTATACAGGAGCTGAAAAATAGTCTTATATAGCCTTGCGAGGATTTGAGCTAGT 660

Qy 661 TTGATTTAATACATGGGCCAACCTAGTGTCTTGGTTTTTGACCTGCTTTCAAAAAAAGT 720

Db 661 TTGATTTAATACATGGGCCAACCTAGTGTCTTGGTTTTTGACCTGCTTTCAAAAAAAGT 720

Qy 721 ATTAATTTTATTTATCTTAGAAATGGGTG 749

Db 721 ATTAATTTTATTTATCTTAGAAATGGGTG 749

## RESULT 2

ID	AA	Size	GC	GC%	GC60%	GC70%	GC80%	GC90%	GC100%	GC110%	GC120%	GC130%	GC140%	GC150%	GC160%	GC170%	GC180%	GC190%	GC200%	GC210%	GC220%	GC230%	GC240%	GC250%	GC260%	GC270%	GC280%	GC290%	GC300%	GC310%	GC320%	GC330%	GC340%	GC350%	GC360%	GC370%	GC380%	GC390%	GC400%	GC410%	GC420%	GC430%	GC440%	GC450%	GC460%	GC470%	GC480%	GC490%	GC500%	GC510%	GC520%	GC530%	GC540%	GC550%	GC560%	GC570%	GC580%	GC590%	GC600%	GC610%	GC620%	GC630%	GC640%	GC650%	GC660%	GC670%	GC680%	GC690%	GC700%	GC710%	GC720%	GC730%	GC740%	GC750%	GC760%	GC770%	GC780%	GC790%	GC800%	GC810%	GC820%	GC830%	GC840%	GC850%	GC860%	GC870%	GC880%	GC890%	GC900%	GC910%	GC920%	GC930%	GC940%	GC950%	GC960%	GC970%	GC980%	GC990%	GC1000%	GC1010%	GC1020%	GC1030%	GC1040%	GC1050%	GC1060%	GC1070%	GC1080%	GC1090%	GC1100%	GC1110%	GC1120%	GC1130%	GC1140%	GC1150%	GC1160%	GC1170%	GC1180%	GC1190%	GC1200%	GC1210%	GC1220%	GC1230%	GC1240%	GC1250%	GC1260%	GC1270%	GC1280%	GC1290%	GC1300%	GC1310%	GC1320%	GC1330%	GC1340%	GC1350%	GC1360%	GC1370%	GC1380%	GC1390%	GC1400%	GC1410%	GC1420%	GC1430%	GC1440%	GC1450%	GC1460%	GC1470%	GC1480%	GC1490%	GC1500%	GC1510%	GC1520%	GC1530%	GC1540%	GC1550%	GC1560%	GC1570%	GC1580%	GC1590%	GC1600%	GC1610%	GC1620%	GC1630%	GC1640%	GC1650%	GC1660%	GC1670%	GC1680%	GC1690%	GC1700%	GC1710%	GC1720%	GC1730%	GC1740%	GC1750%	GC1760%	GC1770%	GC1780%	GC1790%	GC1800%	GC1810%	GC1820%	GC1830%	GC1840%	GC1850%	GC1860%	GC1870%	GC1880%	GC1890%	GC1900%	GC1910%	GC1920%	GC1930%	GC1940%	GC1950%	GC1960%	GC1970%	GC1980%	GC1990%	GC2000%	GC2010%	GC2020%	GC2030%	GC2040%	GC2050%	GC2060%	GC2070%	GC2080%	GC2090%	GC2100%	GC2110%	GC2120%	GC2130%	GC2140%	GC2150%	GC2160%	GC2170%	GC2180%	GC2190%	GC2200%	GC2210%	GC2220%	GC2230%	GC2240%	GC2250%	GC2260%	GC2270%	GC2280%	GC2290%	GC2300%	GC2310%	GC2320%	GC2330%	GC2340%	GC2350%	GC2360%	GC2370%	GC2380%	GC2390%	GC2400%	GC2410%	GC2420%	GC2430%	GC2440%	GC2450%	GC2460%	GC2470%	GC2480%	GC2490%	GC2500%	GC2510%	GC2520%	GC2530%	GC2540%	GC2550%	GC2560%	GC2570%	GC2580%	GC2590%	GC2600%	GC2610%	GC2620%	GC2630%	GC2640%	GC2650%	GC2660%	GC2670%	GC2680%	GC2690%	GC2700%	GC2710%	GC2720%	GC2730%	GC2740%	GC2750%	GC2760%	GC2770%	GC2780%	GC2790%	GC2800%	GC2810%	GC2820%	GC2830%	GC2840%	GC2850%	GC2860%	GC2870%	GC2880%	GC2890%	GC2900%	GC2910%	GC2920%	GC2930%	GC2940%	GC2950%	GC2960%	GC2970%	GC2980%	GC2990%	GC3000%	GC3010%	GC3020%	GC3030%	GC3040%	GC3050%	GC3060%	GC3070%	GC3080%	GC3090%	GC3100%	GC3110%	GC3120%	GC3130%	GC3140%	GC3150%	GC3160%	GC3170%	GC3180%	GC3190%	GC3200%	GC3210%	GC3220%	GC3230%	GC3240%	GC3250%	GC3260%	GC3270%	GC3280%	GC3290%	GC3300%	GC3310%	GC3320%	GC3330%	GC3340%	GC3350%	GC3360%	GC3370%	GC3380%	GC3390%	GC3400%	GC3410%	GC3420%	GC3430%	GC3440%	GC3450%	GC3460%	GC3470%	GC3480%	GC3490%	GC3500%	GC3510%	GC3520%	GC3530%	GC3540%	GC3550%	GC3560%	GC3570%	GC3580%	GC3590%	GC3600%	GC3610%	GC3620%	GC3630%	GC3640%	GC3650%	GC3660%	GC3670%	GC3680%	GC3690%	GC3700%	GC3710%	GC3720%	GC3730%	GC3740%	GC3750%	GC3760%	GC3770%	GC3780%	GC3790%	GC3800%	GC3810%	GC3820%
----	----	------	----	-----	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------

AC AAX30099

DT 17-JUN-1999 (first entry)

DE Borrellia afzelli ACAI protein encoding cDNA

KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss

OS Borrelia afzelii

FH	Key	Location/Qualifiers

/\*tag=

PN W09912960-A2

PD 18-MAR-1999.  
XX

PF 04-SEP-1998; 98WO-IB01424.  
YY

PR	16-SEP-1997;	97US-0059036.
PR	10-SEP-1997;	97DK-0001041

XX (SYMB-) SYMBITCOM AB.  
PA

AA Berastrom S.  
PT

AA WPT: 1999-215  
DR

DK F-PSDB; AAU042/9.  
XX

PT protein P13

xx  
ps Claim 1: Page 110-111: 118bn: English

The present sequence encodes a Borrel

CC present/invention describes an isolated nucleic acid (I) that: (i)  
CC encodes a polypeptide fragment (II) immunologically reactive with  
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of  
CC *Borrelia burgdorferi* B313, but not reactive with most proteins from  
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions  
CC to specific nucleic acid sequences but not to genomic DNA from most  
CC other spirochaetes. (I), (II) and transformed cells are useful in  
CC vaccines to protect against infection by *B. burgdorferi* sensu lato.  
CC (I), (II) and antibodies raised against (II) are used to diagnose such  
CC infections, in standard immunoassays or amplification/hybridization  
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein  
CC is involved in virulence and is highly conserved within *B. burgdorferi*  
CC sensu lato, but is absent from other *Borrelia* species (e.g. those  
CC responsible for relapsing fever or yaws borreliosis). It should provide  
CC a more specific and sensitive antibody response, and diagnosis,  
CC compared with use of whole bacteria as antigen.

50 Sequence 862 BP; 265 A; 94 C; 162 G; 341 T; 0 other;

Query Match	81.08;	Score 607;	DB 20;	Length 862;
-------------	--------	------------	--------	-------------

Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

```
OY 1 TAGAATTTTCAACAATTAAGATATTGTTAAAGAATTGAATTGCTAATTTATGTTA 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27 TAGAATTTTTCGACAAATAAAGACATATTATTAAGAATTGAATTGCTAATTTATGTTA 86
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 AATCAAGAAGCTCTATGTGTAAGCAAAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 87 AATCAAGAAGCTCTATGTGTAAGCAAAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 121 AATAGTTCACAAAGCCTTTTAAATTTTCATTAATATGCTACCATTAATACGAG-TTTAAT 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 147 AATAGTTCACAAAGCCTTTTAAATTTTCATTAATATGCTACCATTAATACGAGTTTAAAT 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 180 AAGAGGTTTTTATGTAATAGTTAATTTTATTTGTTGCTATGCTATGCTATGCTATGCT 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 207 AAGAGGTTTTTATGTAATAGTTAATTTTATTTGTTGCTATGCTATGCTATGCTATGCT 266
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 240 AGTTTTCCTCAA--GATGATTTCTAAAGCACTTTTAACTCTGGAGCGGAGAAAAAATT 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 267 AGCTTTGCTCAAGCTGATGATTTCTAAAGCGCTTTTAAATTTGAGAGCGGAGAAAAAATT 326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 297 TTGGTTTATGAAGCTTAATAGAAAGATTCCTTGTACCATTTTATGTAACCTTTTATTA 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 327 TTAGCTTATGAAGCTTAATAGAAAGATTCCTATGTCCTATTTTATGTAACCTTTTATTA 386
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 357 GGGTTCGGGATGATGTTCTTTTGTCTCAAGAGATATCCTTGGAGGTTCTTATTTCTTGA 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 387 GGGTTCGGGATGATGTTCTTTTGTCTCAAGAGATATCCTTGGAGGTTCTTATTTCTTGA 446
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 417 TTGATGCGGTTGATATAGGTTAATACTTACAGAGCTTAAATTTGACATCAAGATTTT 476
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 447 TTGATGCGGTTGATATAGGTTAATACTTACAGAGCTTAAATTTGACATCAAGATTTT 506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 477 GATTAATATGCTAAAGAAAGCTGATTTTAAAGTGAAGCTTGGGTTAAGGAAATGATGTGCA 536
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 507 GATTAATATGCTAAAGAAAGCTGATTTTAAAGTGAAGCTTGGGTTAAGGAAATGATGTGCA 566
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 537 GGTGGTTCATGATGCGTGTGACAGAGATGACAGAAATGCTTCCATTACATTTGCT 596
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 567 GGTGGTTCATGATGCGTGTGACAGAGATGACAGAAATGCTTCCATTACATTTGCT 626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 597 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 656
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 627 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 657 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 716
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 687 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 717 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 743
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 747 AATTAATTAACAGAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 773
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 3  
AAK30098  
ID AAK30098 standard; DNA; 759 BP.

AAK30098;

17-JUN-1999 (first entry) .

Borrelia burgdorferi B31 protein encoding cDNA.

Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.

Borrelia burgdorferi.

Key Location/Qualifiers

FT CDS 170..709 /tag= a

PN W09912960-A2.

```
XX 18-MAR-1999.
PD 04-SEP-1998; 98WO-1B01424.
XX
PE 16-SEP-1997; 97US-0059036.
XX
PR 10-SEP-1997; 97DK-0001041.
XX
PA (SYMB) SYMBICOM AB.
XX
PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
XX
PT P-PSDB; AAV04278.
XX
PT Nucleic acid from Borrelia burgdorferi encoding virulence associated
XX protein P13
XX
XX Claim 1; Page 107-108; 118pp; English.
XX
CC The present sequence encodes a Borrelia burgdorferi B31 protein. The
CC present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by B. burgdorferi sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections. In standard immunoassays or amplification/hybridization
CC tests, (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within B. burgdorferi
CC sensu lato, but is absent from other Borrelia species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.
XX
SQ Sequence 759 BP; 228 A; 90 C; 157 G; 284 T; 0 other;
XX
Query Match 74.5%; Score 558.2; DB 20; Length 759;
Best Local Similarity 87.0%; Pred. No. 6; 3e-98;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;
```

```

Db      421 GCTTATCTCTGCGGGGCTTATTTGGATATCAAGCGCTGATGTTACTTAAAGC 480
Qy      497 TGATTTTAAGGACTTGGGTAAGGAATGATGTGACAGTGTGTTACTATAGGCTGT 556
Db      481 TGCCTTTCAATGACTTGGGTAAGGAGTATGTATGACAGTGTGTTACTATAGGCTGT 540
Qy      557 GACAAAGATTGACAGAAATGTTCTTCATTTACATTTGCTAATATTAATTAACAGAACT 616
Db      541 GACAAAGATTGACAGAAATTTATTTCTTCATTTACATTTGCTAATATTAATTAACAGAACT 600
Qy      617 GAAAAATAGTCTTAAATTAAGCTTGGGAGATTTGAGCTGTTGGTATTAACATGCG 676
Db      601 AAAAAATAGTCTTAAATTAAGCTTGGGAGATTTGAGCTGTTGGTATTAACATGCG 660
Qy      677 CCAAGTAGTCTCTGTTGGTGGTGGAGTGTCTTCAAAAAAGCTATTAATTTATTC 736
Db      661 CCATTCAGTCTCTTGGGTTTGAAGTGTCTTCAAAAAAGCTATTAATTTATTC 720
Qy      737 TAGAAATGGGTG 749
Db      721 ACAAAATGGGTG 733

```

## RESULT 4

AAK61505 standard; DNA; 540 BP.

AAK61505;

19-JUL-1999 (first entry)

B. burgdorferi antigenic protein coding sequence, f933.nt.

Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

Borrelia burgdorferi.

MO9859071-A1.

30-DEC-1998.

18-JUN-1998; 98WO-US12718.

03-SEP-1997; 97US-0057483.

20-JUN-1997; 97US-0050359.

22-JUL-1997; 97US-0053344.

22-JUL-1997; 97US-0053377.

(HUMA-) HUMAN GENOME SCI INC.

(MEDI-) MEDIMUNE INC.

Choi GH, Erwin AL, Hanson MS, Lathigra R;

WPI; 1999-189980/16.

P-PSDB; AAY19808.

New isolated Borrelia burgdorferi nucleic acids - used to develop

products for the diagnosis, prevention and treatment of diseases

caused by Borrelia, particularly Lyme disease

Claim 1; Page 73; 275pp; English.

This sequence encodes a Borrelia burgdorferi (Bb) protein of the invention, which is suitable for use in a vaccine. The Bb polypeptides can be used in vaccines for eliciting protective antibodies to members of the Borrelia genus, particularly for the use against Lyme disease in humans and animals. They can be used for preventing or attenuating an infection caused by a member of the Borrelia genus. The products can also be used for detection of members of the Borrelia genus.

Query Match 53.8%; Score 402.8; DB 20; Length 540;  
Best Local Similarity 85.6%; Pred. No. 3.1e-68;  
Matches 462; Conservative 0; Mismatches 72; Indels 6; Gaps 1;

```

Qy      192 ATGATAGTATTTTAAATTTTATTTGTTGTAATCTTTTGTGCTTTTCTAGTTGCAAA 251
Db      1  ATGATTAATCTTTTAAATTTTGTGTTGTAATCTTTTGTGCTTTTCTAGTTGCTCAA 60
Qy      252 GATGATCT-----AAAGACCTTTTAAATCTGGAGCGGAGAAAAATTTTGGTTAT 305
Db      61  GCTATGATTTCTAAATATGTCGCTTTGGATGAGTGTGTGAGAAAACTTTTGTGTTAT 120
Qy      306 GAACTAATTAAGAAAGATCTCTTGTACATTTTATTTGAACCTTTTATAGGTTGGG 365
Db      121 GAACTAGCAAGCAAGATCTTATTTGACATTTTATTTGAACCTTTTATAGGTTGGA 180
Qy      366 ATAGGCTCTTTTGTCAAGAGATATCCCTTGAGGTTCTCTTATCTTGATTTGATGCG 425
Db      181 ATAGGCTCTTTGCTCAAGGAGATATTTCTTGAGGTTCTCTTATCTTGATTTGATGCG 240
Qy      426 GTTGATATAGGTTAATATCTTACAGAGCTTATTTGACATCAAGATTTTGATATAT 485
Db      241 GTTGATATAGGCTTATATCTTACAGAGCTTATTTGATCAAGAGCTTATGATATAT 300
Qy      486 GCTAAAAAGCTGATTTTAACTGAGCTGGGTAAGGAATGATGTGGCGATGCTGTT 545
Db      301 ACTAAAAAGCTGCTTTTCAATGAGCTGGGTAAGGATGATGTAAGGCTGCTGTT 360
Qy      546 ACTATGCTGTGACAGATGACAGAAATGTTCTTCATTTACATTTGCTAATATATAT 605
Db      361 ACTATGCTGTGACAGATGACAGAAATTTCTTCATTTACATTTGCTAATATATAT 420
Qy      606 AACAGAGAGCTGAAAAATATGCTTAATATATGCTTGGAGAGATTTGAGCTTATGAT 665
Db      421 AATAGGAAGCTAAAAAATATGCTTAATATATGCTTGGAGAGATTTGAGCTTATGAT 480
Qy      666 ATTAACATGGGCCCAAGCTAGTGTCTTGTGTTGAGAGTCTTTCAAAAAAGCTATTA 725
Db      481 GTTGCAATGGGCCAATCCAGTGTCTTGTGTTGAACTGCTTTCAAAAAAGCTATTA 540

```

## RESULT 5

AAK61506 standard; DNA; 477 BP.

AAK61506;

19-JUL-1999 (first entry)

B. burgdorferi antigenic protein coding sequence, f933.nt.

Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

Borrelia burgdorferi.

MO9859071-A1.

30-DEC-1998.

18-JUN-1998; 98WO-US12718.

03-SEP-1997; 97US-0057483.

20-JUN-1997; 97US-0050359.

22-JUL-1997; 97US-0053344.

22-JUL-1997; 97US-0053377.

(HUMA-) HUMAN GENOME SCI INC.

(MEDI-) MEDIMUNE INC.

Choi GH, Erwin AL, Hanson MS, Lathigra R;

WPI; 1999-189980/16.

P-PSDB; AAY19808.



XX New isolated Borrelia burgdorferi nucleic acids - used to develop  
PT products for the diagnosis, prevention and treatment of diseases  
PT caused by Borrelia, particularly Lyme disease  
PS  
XX  
XX Claim 1; Page 73; 275pp; English.

CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the  
CC invention, which is suitable for use in a vaccine. The Bb polypeptides  
CC can be used in vaccines for eliciting protective antibodies to members of  
CC the Borrelia genus, particularly for the use against Lyme disease in  
CC humans and animals. They can be used for preventing or attenuating an  
CC infection caused by a member of the Borrelia genus. The products can also  
CC be used for detection of members of the Borrelia genus.

XX Sequence 477 BP; 132 A; 62 C; 114 G; 169 T; 0 other;

Query Match 49.2%; Score 368.8; DB 20; Length 477;  
Best Local Similarity 85.9%; Pred. No. 9,7e-62;  
Matches 409; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

250 AAGATGATCTTAAAGACACCTTTTAACTGGAGCGGAGAAAAATTTGGTTATGAAA 309  
1  
2 ATGATCTTAAATGCGTGGTGGATGAGTGGTGAAGAAAACTTTGGTTATGAAA 61  
310 CTAATAGAAAGATCTCTTGACATTTTATGAACTTTTATAGGCTCGGATAG 369  
1  
62 CTAAGCAAGCAAGATCTTATGACATTTTATGAACTTTTATAGGCTCGGATAG 121  
370 GTTCTTTCTCAAGAGATCTCTGAGGTTCTTATCTTGGATTGATGGGTTG 429  
1  
122 GCTCCTTCTCAGAGATATCTTGGAGTTCTTATCTTGGATTGATGGGTTG 181  
430 GTATAGGTTAACTTACAGAGCTTATTTGACATCAAGATTTGATTAATATGCTA 489  
1  
182 GTATAGGCTTACTTACCTGCGGGCTTATTTGATCAAGGCTTATGATGATGCTA 241  
490 AAAAGCTATTTTAAAGCTGAGGTTGAGGATGATGTTGCGAGGTGCTGATGCTA 549  
1  
242 AAAAGCTATTTTAAAGCTGAGGTTGAGGATGATGTTGCGAGGTGCTGATGCTA 301  
550 TGGCTGTGACAGATGACAGAAATGCTTCCATTTACATTTGCTAATAATTATACA 609  
1  
302 TGGCTGTGACAGATGACAGAAATGCTTCCATTTACATTTGCTAATAATTATACA 361  
610 GGAAGCTGAAAAATGCTTAAATATAGCCTTGGAGGATTTGAGCTAGTTGATATTA 669  
1  
362 GGAAGCTGAAAAATGCTTAAATATAGCCTTGGAGGATTTGAGCTAGTTGATATTA 725  
670 ACATGGGCAAGCTGCTGCTTGGTTTGAAGCTGCTTCAAAAAAACCTATTA 725  
1  
422 CAATGGGCAAGCTGCTGCTTGGTTTGAAGCTGCTTCAAAAAAACCTATTA 477

RESULT 6  
AAK20310/C  
ID AAK20310 standard; DNA; 1761 BP.

AAK20310;  
04-MAY-1999 (first entry)

DE Borrelia burgdorferi polynucleotide sequence #63.

KM Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;  
KM epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
KM infection; diagnosis; characterisation; detection; ds.

OS Borrelia burgdorferi.

XX W09858943-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12764.  
XX  
XX 03-SEP-1997; 97US-0057483.  
XX  
XX 20-JUN-1997; 97US-0050359.  
XX  
XX 22-JUL-1997; 97US-0053344.  
XX  
XX 22-JUL-1997; 97US-0053377.

PA (HUMAN-) HUMAN GENOME SCI INC.  
PA (MEDI-) MEDIMUNE INC.

PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;

PI White OR;

DR WPI; 1999-081217/07.

PT New isolated Borrelia burgdorferi nucleic acids - used to develop  
PT products for the detection, diagnosis, characterisation, prevention  
PT and therapy of infections, particularly Lyme disease

PS Claim 1; Page 1045-1046; 1128pp; English.

AAK20248 to AAK20402 represent polynucleotide sequences isolated from  
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for  
CC the detection, diagnosis, characterisation, prevention and therapy of  
CC Bb infections, e.g. Lyme disease. They can also be used for the  
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs  
CC to a family of motile, spiral-shaped bacteria called Spirochetes.  
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and  
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as  
CC Lyme disease.

SQ Sequence 1761 BP; 548 A; 267 C; 268 G; 676 T; 2 other;

Query Match 13.1%; Score 98.4; DB 20; Length 1761;  
Best Local Similarity 49.7%; Pred. No. 5.2e-10;  
Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;

1 TAGAATTTTCAACAAATTAAGATATTTGTTAAAGAAATGGAATTTGCTAATTTATGCTTA 60  
1  
1031 TGGCATACAAAAAATAAGATAGCATTAATTTCAAAATTTAAAGAGTTTATGAAAACTTT 972  
61 AATCAGAAGCTCTATTTGTAAGCGAATTTCCAGTACAACTTGAAAAAGTTAATTA 120  
1  
971 ACATTAATAATCTGTAGAACACGAAATTCACAAATATCTTTAGAAAAAGTTGA 912  
121 AATAGTTCCAAAAAGCTTTTAAATTTCAATTAATGCTACATTAATACCAAGTTAATA 180  
1  
911 ACTCCATTAATTAATCTAAAAATTTGCTACATTTGTTTATCAAAACTTAATTTAG 852  
181 AAGG--GGTTTATGAATTAAGTTTAAATTTTAAATTTTGAATCTTTTGTGCTTTT 237  
1  
851 GAGGTAATTAATTAAGAAAAAATTTTCAATTAATTAATTTTAAATTTTAAATTAAC 792  
238 CTAGTTTGCCTAAGATGATTTCAAAAGCACTTTTAAATCGAGCGGAGGAAAAATTTT 297  
1  
791 AATCTTTGCACAGATTAACCTTGAANAAGTGTGAGT-GGAGATATCGGACCTGTA 733  
298 TGGTTATGAATAATAAGAAAGATTTCTTGTACCATTTTATTAAGACCTTTTATAG 357  
1  
732 TGAATATGAAGCGAAAAAGCAACATCTAGCACCATTTACTTTGAATATCTTTTGT 673  
358 GGTTCGGATAGGTTCTTTTGTCTCAAGAGATATCTTGGAGGTTCTTATTTCTGGAT 417  
1  
672 CTTTACGGATAGGATCTTTTGTCAAGAGATTAATTTGTGTGTGCTGACACTTGCT 613  
418 TTGATCGGTTGGTATAGGTTAATCTTACAGAGCTTATTTGGACATCAAGAGATTTG 477  
1  
612 CTCAGTGTCTGGAAGG-----AATCTTATATGCTGACACATGATCGTAGAGGTA 558  
478 ATAATATGCTAAAAAGCTGATTTTAAGTGCATCTGGGTAAGGAAAGATGTTGCGAG 537  
1  
557 TTGAGGCGTTTACAGAAACACAGCAAGTATATAACCGGAGAGATTAATACAGGAATAG 498

12 ACAATAAAGATATTGTTAAAGAATTGAAATTGCTAATTTATGTTAAATCAGAAGC 71

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053377.



Query Match 8.8%; Score 66; DB 22; Length 7110;  
Best Local Similarity 51.0%; Pred. No. 0.0009;  
Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 35 AATTGAATTCCTATTTATGTTAATCAAGAGCTCATTTGTAGGCAATTCGAG 94  
DB 1672 AATTGAATTTTAAAGATATTTATTTGAAATTTTGTAGATGGGTTAGTTA 1731  
QY 95 TAACAATTTGAAAAAGTTAAATTAATAGTCCAAAGCCCTTTTAAATTCATTAA 154  
DB 1732 TTGGGTAGAGAAAGAAAGAAATTAATAAAGATGATATAGATTTTATTTT 1791  
QY 155 TATGCTACCATATATACAGTTTATTAAGGGTTTATGATAGTTTATTTATTTTAT 214  
DB 1792 AGAATATATTTAATATATAGATATGAGAGGATTTTATTTATTAATTTATATAT 1851  
QY 215 TTGTGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATCTAAAGCATTGTA 274  
DB 1852 TTGATGTTGTTGGTGGTGTGTTAAATTTGTTAATATGATGATTTTATTTATTTA 1911  
QY 275 TCTGGAGCGGAGAAAAATTTTGTATTAGAACTAATATAGAAAGATTCCTGTAC 334  
DB 1912 TTAGAGAGAAATAGAGGATTTGATGTTTAGAGAGTAGTAATAATTTATTTTATTTT 1971  
QY 335 ATTTT 340  
DB 1972 TTTT 1977

RESULT 10  
ABL70587  
ID ABL70587 standard; DNA: 7110 BP.  
AC ABL70587;  
XX 01-JUL-2002 (first entry)  
XX Chemically treated cell signalling DNA sequence#239.  
XX  
XX Cell signalling; cytosine methylation; cell signalling disease;  
XX cancer; tumour; cytostatic; ds.  
XX  
XX Unidentified.  
XX  
XX WO200202807-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 29-JUN-2001; 2001WO-EP07471.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-154758/20.  
XX  
XX Nucleic acid, useful for diagnosis and therapy of diseases associated  
XX with cell signalling e.g. cancer, comprises chemically modified genomic  
XX sequences of genes associated with cell signalling -  
XX  
XX Claim 1; SEQ ID NO 477; 24pp+sequence listing; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of at least  
XX 18 bases of a segment of chemically pretreated DNA of genes associated  
XX with cell signalling. The activity of the modified sequences of the  
XX invention may be described as cytostatic. The object of the invention is  
XX to provide the chemically modified DNA of genes associated with cell  
XX signalling, as well as oligonucleotides and/or PNA-oligomers for  
XX detecting cytosine methylations, as well as a method which is

CC particularly suitable for the diagnosis and/or therapy of genetic and  
CC epigenetic parameters of genes associated with cell signalling. The  
CC chemically modified DNA provided by the invention is useful for diagnosis  
CC and therapy of diseases such as solid tumours and cancer. The sequences  
CC given in records ABL70111-ABL70626 represent chemically pre-treated  
CC genomic DNA's of genes associated with cell signalling.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification, but is based on sequence information supplied by the  
CC European Patent Office.  
XX  
XX Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;  
SQ

Query Match 8.8%; Score 66; DB 24; Length 7110;  
Best Local Similarity 51.0%; Pred. No. 0.0009;  
Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 35 AATTGAATTCCTATTTATGTTAATCAAGAGCTCATTTGTAGGCAATTCGAG 94  
DB 1672 AATTGAATTTTAAAGATATTTATTTGAAATTTTGTAGATGGGTTAGTTA 1731  
QY 95 TAACAATTTGAAAAAGTTAAATTAATAGTCCAAAGCCCTTTTAAATTCATTAA 154  
DB 1732 TTGGGTAGAGAAAGAAAGAAATTAATAAAGATGATATAGATTTTATTTT 1791  
QY 155 TATGCTACCATATATACAGTTTATTAAGGGTTTATGATAGTTTATTTATTTTAT 214  
DB 1792 AGAATATATTTAATATATAGATATGAGAGGATTTTATTTATTAATTTATAT 1851  
QY 215 TTGTGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATCTAAAGCATTGTA 274  
DB 1852 TTGATGTTGTTGGTGGTGTGTTAAATTTGTTAATATGATGATTTTATTTATTTA 1911  
QY 275 TCTGGAGCGGAGAAAAATTTTGTATTAGAACTAATATAGAAAGATTCCTGTAC 334  
DB 1912 TTAGAGAGAAATAGAGGATTTGATGTTTAGAGAGTAGTAATAATTTATTTTATTTT 1971  
QY 335 ATTTT 340  
DB 1972 TTTT 1977

RESULT 11  
ABL33556  
ID ABL33556 standard; DNA: 7110 BP.  
AC ABL33556;  
XX 26-MAR-2002 (first entry)  
XX  
XX Human immune system associated gene SEQ ID NO: 1529.  
XX  
XX Human immune system associated gene  
XX  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
XX antiarteriosclerotic; antinaemic; cytostatic; noctropic;  
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
XX antitumour; cancer; eye disease; arteriosclerosis; anaemia;  
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
XX gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200200928-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX

PI Olek A, Plepenbrock C, Berlin K;  
 XX WPI: 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 XX Claim 1: SEQ ID NO 1529; 32pp + Sequence Listing; German.  
 XX  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC leukodystrophy, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX  
 XX Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;  
 SQ  
 Query Match 8.8%; Score 66; DB 24; Length 7110;  
 Best Local Similarity 51.0%; Pred. No. 0.0009;  
 Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;  
 OY 35 AATTGAATTGCTAATTTATGCTTAATCAAGAGCTATATGGTAACGATTTCCGAG 94  
 Db 1672 AATTGAATTGCTAATTTATGCTTAATCAAGAGCTATATGGTAACGATTTCCGAG 1731  
 OY 95 TAACAATTTGAAAAAGTTAAATTAATGATTCGAAAGCCCTTTTAAATTCATTTAA 154  
 Db 1732 TTTGGGTAGAAAAGGAAAGAAATTAATTAATGATTCGAAAGCCCTTTTAAATTCATTTAA 1791  
 OY 155 TATGCTACCAATATACCACTTAATTAAGGGGTTTATGAAATAGCTTTTAAATTTTAA 214  
 Db 1792 AGAAATTAATTTAATTAATGAGATAGAGAGATTTTATTTAATTTTAAATTAATTAAT 1851  
 OY 215 TTTGGTAATCTTTTGTGCTTTTCTAGTTTTCGTCAGATGATTCGAAAGCACTTTAA 274  
 Db 1852 TTGATGTTGTTTGTGGTGGTTAAATTAATTTTAAATGATTTTAAATTTTAAATTA 1911  
 OY 275 TCTGGAGCGGAGAAAAATTTTGGTTATGAACTAATTAAGAAATTCCTTGTACC 334  
 Db 1912 TTAGAGAGAGATGAGAGGATTTGATGTTTGAAGAAATTAATTTTAAATTTTAA 1971  
 OY 335 ATTTT 340  
 Db 1972 TTTT 1977  
 ASUFT 12  
 AAS61290  
 ID AAS61290 standard; DNA; 7110 BP.  
 XX  
 XX AAS61290;  
 AC  
 XX 29-JAN-2002 (first entry)  
 DT  
 XX  
 XX Human gene regulation-associated gene oligonucleotide #245.  
 DE  
 XX  
 XX Human; Gene regulation-associated gene; severe combined immunodeficiency;  
 KM cardiac damage; inflammatory response; Hemophilia; Werner syndrome;  
 KM asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;  
 KM renal disease; Preclampsia; cardiac allograft vascular disease;  
 KM colorectal cancer; thyroid cancer; oesophageal cancer; ds. tumour;  
 KM immunostimulant; cardiac; anti-inflammatory; coagulant; antisthmatic;  
 KM nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200177375-A2.  
 PN  
 XX  
 XX 18-OCT-2001.

XX  
 PF 06-APR-2001; 2001WO-EP03968.  
 XX  
 PR 06-APR-2000; 2000DE-1019058.  
 XX  
 PR 07-APR-2000; 2000DE-1019173.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.  
 PA  
 XX  
 XX Olek A, Plepenbrock C, Berlin K;  
 PI  
 XX WPI: 2002-017470/02.  
 DR  
 XX  
 XX New nucleic acid sequences from chemically modified genes associated  
 PT with gene regulation, useful for analysing cytosine methylations for  
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency  
 PT disease -  
 XX  
 XX Disclosure: SEQ ID NO 251; 26pp; English.  
 PS  
 XX  
 XX The invention relates to 224 nucleic acid sequences comprising at least  
 CC 18 bases of a chemically pretreated gene associated with gene regulation  
 CC selected from 43 known genes (or complementary sequences). The  
 CC chemical pretreatment converts cytosine bases unmethylated at the  
 CC 5-position to uracil or another base with hybridisation behaviour  
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.  
 CC The DNA sequences, oligomers (or sets/arrays) and method are  
 CC useful in the diagnosis of diseases (or predisposition to diseases)  
 CC associated with gene regulation and in therapy of such diseases, by  
 CC enabling analysis of the cytosine methylation patterns of such genes,  
 CC kits are provided. They are especially useful in diagnosis  
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac  
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,  
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,  
 CC preclampsia, graft versus-host disease. The present sequence is a  
 CC sequence included in the sequence data for this specification and is  
 CC associated with the human gene regulation-associated genes.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 SQ  
 Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;  
 Query Match 8.8%; Score 66; DB 24; Length 7110;  
 Best Local Similarity 51.0%; Pred. No. 0.0009;  
 Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;  
 OY 35 AATTGAATTGCTAATTTATGCTTAATCAAGAGCTATATGGTAACGATTTCCGAG 94  
 Db 1672 AATTGAATTGCTAATTTATGCTTAATCAAGAGCTATATGGTAACGATTTCCGAG 1731  
 OY 95 TAACAATTTGAAAAAGTTAAATTAATGATTCGAAAGCCCTTTTAAATTCATTTAA 154  
 Db 1732 TTTGGGTAGAAAAGGAAAGAAATTAATTAATGATTCGAAAGCCCTTTTAAATTCATTTAA 1791  
 OY 155 TATGCTACCAATATACCACTTAATTAAGGGGTTTATGAAATAGCTTTTAAATTTTAA 214  
 Db 1792 AGAAATTAATTTAATTAATGAGATAGAGAGATTTTATTTAATTTTAAATTAATTAAT 1851  
 OY 215 TTTGGTAATCTTTTGTGCTTTTCTAGTTTTCGTCAGATGATTCGAAAGCACTTTAA 274  
 Db 1852 TTGATGTTGTTTGTGGTGGTTAAATTAATTTTAAATGATTTTAAATTTTAAATTA 1911  
 OY 275 TCTGGAGCGGAGAAAAATTTTGGTTATGAACTAATTAAGAAATTCCTTGTACC 334  
 Db 1912 TTAGAGAGAGATGAGAGGATTTGATGTTTGAAGAAATTAATTTTAAATTTTAA 1971  
 OY 335 ATTTT 340  
 Db 1972 TTTT 1977

ID	AB067059	standard; DNA; 34688 BP.
XX	AB067059;	
XX	28-AUG-2002	(first entry)
XX	Human angiogenesis associated polynucleotide SEQ ID NO 89.	
XX	Human; angiogenesis; methylation; eye disease; glaucoma; tumour;	
XX	inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;	
XX	macular degeneration; inflammatory bowel disease; Crohn's disease;	
XX	antirheumatic; antirheumatic; antidiabetic; antipsoriatic;	
XX	antiartherosclerotic; ds.	
XX	Homo sapiens.	
XX	WO200246454-A2.	
XX	13-JUN-2002.	
XX	06-DEC-2001; 2001WO-EPI4320.	
XX	06-DEC-2000; 2000DE-1061338.	
XX	(EPIG-) EPIGENOMICS AG.	
XX	Schacht O;	
XX	WPI; 2002-500450/53.	
XX	New nucleic acid fragments from chemically treated	
XX	angiogenesis-associated genes, useful for determining methylation	
XX	status; e.g. in diagnosis or treatment of cancer	
XX	Claim 1; SEQ ID NO 89; 41pp + Sequence listing; German.	
XX	The invention relates to a nucleic acid (I) comprising a segment of 18	
XX	bases of chemically pretreated DNA of angiogenesis-associated genes (II)	
XX	having sequences (AB066771-AB067178) or their complements. (1), also	
XX	related oligomers, are used to evaluate the methylation status and/or	
XX	single-nucleotide polymorphisms, in angiogenesis-related genes, for	
XX	diagnosis and treatment of eye diseases, proliferative retinopathy,	
XX	neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,	
XX	diabetic retinopathy, macular degeneration caused by neovascularisation,	
XX	psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and	
XX	Crohn's disease.	
XX	Note: The sequence data for this patent did not form part of the printed	
XX	specification, but was obtained in electronic format directly from WIPO	
XX	at ftp.wipo.int/pub/published_pct_sequences.	
XX	Sequence 34688 BP; 9920 A; 444 C; 6941 G; 17383 T; 0 other;	
XX	Query Match 8.8%; Score 66; DB 24; Length 34688;	
XX	Best Local Similarity 51.0%; Pred. No. 0.001;	
XX	Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0	
XX	35 AATTGAATTTGCTAATTTTATGTTAAATTCAGAACGCTCTATTTGGTAACGAATTTCCGAG 94	
XX	1672 AATTGAATTTTAAAAAAGATTAATTTTTCGAAATTTTTCAGTTAGATGAGGATGCTTA 1731	
XX	95 TAACAATTTGAAAAAAGTTAAATTTTAAATAGTCCAAAAAGCCCTTTTAAATTCATTAA 154	
XX	1732 TTTGGGTGAGAAAGCAAAAGCAAAATAAAAAATTTGGAAAAGGATGATTTATTTT 1791	
XX	155 TATGCTACCATTAATACCAAGTTTAATAAAGGGCTTTTATGAAATTAAGTTTATTTTAA 214	
XX	1792 AGAATTAATTTTAAATTAATTAAGAAATAGAGAGATTTTATTTAATTTTAAATTAAT 1851	
XX	215 TTTGGTAACTTTTGTGCTTTTCTTCAAGTTGATTTCAAAAGCATTTTAA 274	

Db	1852	TTGAGTGTGTTGTTGGTGTAAATATTGTTAAATATGATGATTTTTTATTTATTA	1911
Qy	275	TCCTGGGACGGGAGAAAATTTTGGTTTGTGAACATAAAGAAAGATTCCTGTAC	334
Db	1912	TTAAGAGGAAGTACGAGGATTTGATGTTTACGACGTAAGATTTATTTTATTTT	1971
Qy	335	ATTTT 340 	
Db	1972	TTTTT 1977	
RESULT 14			
ID	ABK40058		
XX	ABK40058	standard; DNA; 7441 BP.	
AC	ABK40058;		
XX			
DT	21-MAY-2002	(first entry)	
XX			
DE	Human	chemically pretreated gene sequence #70 strand 2.	
XX			
KW	Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;		
KW	Cytosatic; ALDH6; CYP11A; CYP11B; CYP3A3; DYPD; EPHX2; OCIN; TXNRD1;		
KW	UG8; MR; pharmacogenomics; SNP; single nucleotide polymorphism.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200202806-A2.		
PD			
XX	10-JAN-2002.		
XX			
FE	29-JUN-2001; 2001WO-EP07470.		
XX			
PR	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043826.		
XX			
PA	(EPIC-) EPIGENOMICS AG.		
XX			
F1	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI: 2002-154757/20.		
XX			
PT	New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,		
PT	useful for detecting cytosine methylation state of genes associated		
PT	with pharmacogenomics and for therapy of diseases e.g. cancer		
XX			
PS	Claim 1; SEQ ID No 140; 24pp; English.		
XX			
CC	The invention relates to a nucleic acid comprising a sequence at		
CC	least 18 bases in length of a segment of the chemically pretreated DNA		
CC	sequences associated with pharmacogenomics according to one of the		
CC	sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1		
CC	(NM_000437), CYP3A3 (NM_000776 and NM_017460), DYPD (NM_000110), EPHX2		
CC	(NM_001979), OCIN (NM_002538), TXNRD1 (NM_003330), UG8 (NM_003360),		
CC	MR (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,		
CC	CC (NM_019899) and their complementary sequences, or a sequence (S1) chosen		
CC	from 87 sequences and their complements. The chemical pretreatment		
CC	is bisulphite treatment to convert cytosines (but not methyl-cytosines)		
CC	into uracils. Also included are an oligomer (II) in particular an		
CC	oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in		
CC	each case at least one base sequence having a length of 9 nucleotides		
CC	which hybridises to or is identical to a chemically pretreated DNA of		
CC	genes associated with pharmacogenomics and their complements, arranged in		
CC	an array for analysing diseases associated with the methylation state		
CC	(CpG) and/or detecting SNPs (single nucleotide polymorphisms)		
CC	of the 87 sequences. The oligomers may also be used as PCR primers.		
CC	The set of 87 nucleic acids and their complements is useful for diagnosis		
CC	and therapy of solid tumours and cancer. The present sequence		
CC	represents one the 87 DNA sequences or its complement.		
CC	Note: The sequence data for this patent did not form part		
CC	of the printed specification, but was obtained in electronic		
CC	format directly from Wipo at		
CC	ftp.wipo.int/pub/published_pct_sequences.		



```

Db 6669 TTTTAAGAGTTATTTCGTTTTATTATAGAGTAGATTTTATTTTAAAGTATT 6728
Qy 374 TTTTGCTCAAGAGATATCCTTGAGGTCCTTATTCTTGATTTGATGCGGTGTAN 433
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6729 TATAGTTTAAAGTTATATTTTTTTTTTTTATTAGTATATTGTTAAATAA 6788
Qy 434 AGGTTAATACCTTACAGAGCTTATTGACATCAAGAGATTTGATATATGCTAATAA 493
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6789 TTTTATTTTAATATATAGTTTTTTTGTAAAGTTAGATGAATTTATTTAGT 6848
Qy 494 AGCTGATTTTAACTGACT 512
    || || || || || || || || || || || || || || || || || || ||
Db 6849 TTAGATTTTAGTTTAT 6867

```

Search completed: March 27, 2003, 01:32:55  
 Job time : 204.602 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1051.13 Seconds  
(without alignments) 11540.380 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaatttcacacaaataa.....attatctagaataatgggtg 749

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estbma:\*  
3: em\_estin:\*  
4: em\_estinu:\*  
5: em\_estov:\*  
6: em\_estp1:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_trod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74.8	10.0	1101	17	CNS00LTX2
2	72.4	9.7	1101	17	CNS00EO7
3	71.8	9.6	1101	17	CNS0039G
4	69.6	9.3	996	17	CNS00PFUH
5	68.2	9.1	1101	17	CNS0100X
6	68	9.1	1225	17	CNS0161D

7	67.8	9.1	1013	17	CNS06RPO	AL412260 T7 end of
8	66.6	8.9	910	17	CNS0168P	AL442826 Anopheles
9	65	8.7	1101	17	CNS00DEVL	AL069706 Drosophila
10	64.8	8.7	1201	17	CNS016EQ	AL106652 Drosophila
11	64.2	8.6	928	17	CNS00DKY	AL071865 Drosophila
12	64.2	8.6	1092	17	CNS020K7	AL175696 Tetradon
13	62.8	8.4	1038	17	CNS01TEL7	AL166660 Tetradon
14	62.4	8.3	1204	17	CNS01ME2	AL106628 Drosophila
15	62	8.3	576	17	CNS035N7	AL228940 Tetradon
16	62	8.3	893	17	BH188563	BH188563 036_O_12-
17	62	8.3	893	17	CNS075JW	AL625502 T3 end of
18	61.6	8.2	1101	17	CNS00LTX2	AL078714 Drosophila
19	61.4	8.2	893	17	CNS013XE	AL103436 Drosophila
20	61.4	8.2	1101	17	CNS0042M	AL055440 Drosophila
21	61.4	8.2	1303	14	BH924677	BH924677 AGENCOURT
22	61.2	8.2	1001	17	CNS0155H	AL105023 Drosophila
23	61	8.1	1101	17	CNS00DEDB	AL062360 Drosophila
24	61	8.1	1101	17	CNS00EVL	AL069706 Drosophila
25	61	8.1	1101	17	CNS01219	AL101595 Drosophila
26	60.4	8.1	865	12	BF264686	BF264686 HV-CEA001
27	60.2	8.0	1101	17	CNS00DT7	AL075293 Drosophila
28	60	8.0	1300	10	BE420576	BE420576 HMO00.D0
29	59.8	8.0	524	17	CNS01U90	AL167541 Tetradon
30	59.8	8.0	1184	17	CNS04P4P	AL300850 Tetradon
31	59.6	8.0	987	17	CNS014P0	AL104456 Drosophila
32	59.6	8.0	994	17	CNS01487	AL169872 Tetradon
33	59.6	8.0	1101	17	CNS017KE	AL108152 Drosophila
34	59.4	7.9	865	17	BH180441	BH180441 016_H_20-
35	59.4	7.9	865	17	CNS07MAN	AL617393 T3 end of
36	59.4	7.9	1101	17	CNS003BD	AL064091 Drosophila
37	59.2	7.9	928	17	CNS00DKY	AL071865 Drosophila
38	59.2	7.9	1201	17	CNS016FX	AL106695 Drosophila
39	59	7.9	734	17	CNS010MP	AL099163 Drosophila
40	59	7.9	876	17	CNS009G1	AL053529 Drosophila
41	59	7.9	1092	17	CNS020K7	AL175696 Tetradon
42	59	7.9	1101	17	CNS016DT	AL106619 Drosophila
43	58.8	7.9	1101	17	CNS017B3	AL107817 Drosophila
44	58.6	7.8	876	17	CNS009G1	AL053529 Drosophila
45	58.6	7.8	914	17	CNS002JY	AL097768 Drosophila

## ALIGNMENTS

RESULT 1 1101 bp DNA linear GSS 14-JUN-1999  
CNS00LTX2 Drosophila melanogaster genome survey sequence TET3 end of BAC:  
LOCUS BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AL078714.1 GI:5102004  
Drosophila melanogaster.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mammosser in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial



[illegible]

QY	361	TCGGGATGAGTTCCTTTTSCCTCAGAGATATCCCTTGAGGTCCTTATTCCTGGATTG	420
DY	740	AGTAGRKKRRKRRKRRKRRDTRFEMDDADDDTARDDRRRRRGGDAGCKKTKRRRRRRD	681
QY	421	ATGCGGTTGCGATAGGCGTTAATCTATACAGAGCGTTATTTGACATCAAGATTGTTGATA	480
DY	680	RATMDRDTADAMWADAMWTTTDTDDTDDMKRDRRRRGARRRRRTTARAAMDWTTKAMDMA	621
QY	481	ATAATGTAAAAAAAGCGTATTTTAAAGTCGACTTGCGGTAAGGGAATGATGTTGGCAGCTG	540
DY	620	KWMDKTRADFRMDRNAADWTDKARADRWMAKARARARARDRARARADRRTTCKKTJTA	561
QY	541	TGGTTTACATGCGGTGTCACAGATTGACAGAAATGCTCTCCATTTTACATTTGCTATA	600
DY	560	TWTTWAAAAAAMWAMWAMWTTTATWTTTWTWTTTWTWTTTWTWTTTWTWTTWAAWMTW	501
QY	601	ATTATTAACGAGGAGCGTAAAAATAGTCTTAATATAGCCCTTGAGAGATTGAGCGTACTT	660
DY	500	AAATTAATAAAAAAAMWATTTTTTTTTTTTTTWTWMTAAWMTWMTWTTTWTWMAATTT	441
QY	661	TTGATATTACAT 673	
DY	440	TTTWTWTTWMAAT 428	
RESULT 4			
CNS00FUD/c			
LOCUS	CNS00FUD	996 bp	DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR31021 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL071063		
VERSION	AL071063.1	GI:4951105	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 996)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .		
FEATURES	source		
	1..996		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone="BACR31021"		
	/clone_id="RPc1-98"		
	/note="end : TET3"		
BASE COUNT	383 a	164 c	81 g 171 t 197 others
ORIGIN			
Query Match	9.3%	Score 69.6;	DB 17; Length 996;
Best Local Similarity	36.2%;	Pred. NO. 0.0061;	

OY		1	TAGAAATTTTCAACAAAATAAAGATTGCTTAAGAATGCCTAATTTAAGCTTA	60
Dd		892	MMMMATTTTTTTWMAAMTTTAARTTWRTWTWANTAMWTATTTTTTTTTTTTTTA	833
OY		61	AATCAGAAGCTCTATTGGTAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTA	120
Dd		832	ATMTTGTMTTTTTTTTTTATTTWATTTWATMMMTMAMAMWTTTAMTTMMTTTATATTAATTAAT	773
OY		121	AATAGTCCAAAACCCTTTTTTAATTTCAATTAATGCCATACCATTAATACAGTTAATA	180
Dd		772	TTTTTTTTTTTTTTTTTKATWAATTAATTTATTTTWTATTTTTTTWWATTTTMMTTTTTTTATT	713
OY		181	AAGGGTTTTTTATGAATAAAGTTTTTAAATTTTAAATTTTGATCTGATCTGCTTTTCTA	240
Dd		712	MTTTTTTTTTTTTTTMMWATTTTTTTTTTWTATTTATTTATMTGTATTKATTTAAACGTTTTW	653
OY		241	GTTTTGCTCAAGATGATTTCTAAAAGCACTTTTAACTCGGAGCGGAGCAAAAATTTTTGG	300
Dd		652	TTTTTTTTTRRTTWTWATWATWAAATTTWTAMTTAKAMTTWKAAATTTTGGAAAMWTTTTT	593
OY		301	TTATGTAAACTAATAAAGAAAGATCTCTGTGACAACTTTTATPAGAACCTTTTTTAAAGGT	360
Dd		592	TTTMAAATAIKKADYWAMTTTWTATKTTTTTTTATTTATTTATTTWTKANYYTTTTTTTTKT	533
OY		361	TCGGGATAGGTTCTTTTGCTCAAGAGAGATATCTTGAGAGTTCCTATTCCTGATTTTG	420
Dd		532	TTTTTTTTTATTTTTTTTCCGTTTTTTTTTTTTTTTTTTTTTTTTTCTCRKTTTPTTTNN	473
OY		421	ATCGGCTTGATATAGGGTTAATCTACAGAG	453
Dd		472	NNDGKRGGGGGGVVVVMNNAAAAAAGGG	440

RESULT 5					
CNS0100X					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence sp6 end of BAC				
	BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL098379				
VERSION	AL098379.1	GI:5609990			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 1101)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :				
JOURNAL	BP 101 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
COMMENT	determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and genevieve Payan. It has been constructed in the vector pBelobAC11.				

```

FEATURES:
    source      Location/Qualifiers
1. 1101
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="BACN03G04"
    /clone_1kb="DrosBAC"
    /plasmid="pBelopBAC11"
    /note="end : Sp6"
195 a      108 c      131 g      161 t      506 others
BASE COUNT
      BrTtIn

```

Query Match	9.1%;	Score 68.2;	DB 17;	Length 1101;
Best Local Similarity	18.3%;	Pred. No. 0.01;		
Matches 109;	Conservative 256;	Mismatches 223;	Indels 8;	Gaps 2;

[illegible]

RESULT 6	
CNS0161D/C	
LOCUS	CNS0161D 1225 bp DNA linear GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C18 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL106171
VERSION	ALI06171.1 GI:5620504
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1225)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <a href="http://www.edgp.ebi.ac.uk/">http://www.edgp.ebi.ac.uk/</a> . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector
COMMENT	

FEATURES	Source	Location/Qualifiers
peb10BAC11.	1..1225	
Source	/organism="Drosophila melanogaster"	
	/db_xref="taxon:7277"	
	/clone="BACN15C18"	
	/clone_11p="DrosBAC"	
	/plasmid="peb10BAC11"	
	/note="end : SP6"	
BASE COUNT	266 a 128 c 38 g 368 t 425 others	
ORIGIN		
Query Match	9.1%; Score 68; DB 17; Length 1225;	
Best Local Similarity	33.4%; Pred. No. 0.011;	
Matches 153; Conservative 94; Mismatches 203; Indels 8; Gaps 1;		
OY	4 AATTTCAACAATTAAGATATTTGTTAAAGAATGAATGCTAATTTATGTTAAT 63	
Db	1225 AATTTCAACAATTAAGATATTTGTTAAAGAATGAATGCTAATTTATGTTAAT 1166	
OY	64 CAAGAGCTCTATTGCTAAGCAATTTGAGTACAAATTTGAAAAAGTAAATTAAT 123	
Db	1165 TAT 1106	
OY	124 AGTTCCAAAGAGCTTTTAAATTTCAATTAATATGCTACATTAATACAGTTAATAG 183	
Db	1105 WAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 1054	
OY	184 GGGTTTTATGCAATTAAGTTTAAATTTTATTTTATTTTGGTATCTTTGCTTTTCTAGT 243	
Db	1053 ATWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 994	
OY	244 TTGCTCAAGATGATCTTAAAGCAGCTTTTAAATCTGGAGCGGAGAAAAATTTTGTT 303	
Db	993 WATAAAMWTTAAAAAATAATTAATTTTAAWAAWTTTTRAAAWWTTTTTTWT 934	
OY	304 ATGAACATTAATGAAGAAATCTCTGTCAGCAATTTTATGACCTTTTATAGGCTCG 363	
Db	933 TTTAAATATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 874	
OY	364 GGAATGATCTTTTCTGCTCAAGAGATATCTTGAGGTTCTTATCTTGATTTGATG 423	
Db	873 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 814	
OY	424 CGTTGGATGAGCTTAACTTACAGAGCTTAATTTG 461	
Db	813 CGATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 776	
RESULT 7		
LOCUS	peb10BAC11	
DEFINITION	CNS06RPQ 1013 bp DNA linear GSS 05-JUL-2001	
ACCESSION	AL412260	
VERSION	AL412260.1 GI:12182622	
KEYWORDS	GSS.	
SOURCE	Yarrowia lipolytica.	
ORGANISM	Yarrowia lipolytica	
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Dipodascaceae; Yarrowia.	
AUTHORS	1 (bases 1 to 1013) Souciet,J.L., Aigle,M., Attiguenave,F., Blandin,G., Boitron-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S., deMontigny,J., Dujon,B., Durrens,P., Lepingle,A., Lorente,B., Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,P., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,M., and Weissenbach,J.	
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies	
JOURNAL	FEMS Lett. 487 (1), 3-12 (2000)	
MEDLINE	20584711	
PUBMED	11152876	

AUTHORS	2 (bases 1 to 1013)					
TITLE	Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynrol,C., Attiguenave,F., Wincker,P. and Gallatdin,C.					
JOURNAL	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica					
MEDLINE	FEBS Lett. 487 (1), 95-100 (2000)					
PUBMED	20584727 11552892					
REFERENCE	3 (bases 1 to 1013)					
AUTHORS	Genoscope.					
TITLE	Direct Submission					
JOURNAL	Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)					
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.					
FEATURES	location/Qualifiers					
SOURCE	1..1013 /organism="Yarrowia lipolytica" /strain="CLIB 89" /db_xref="taxon:4952" /clone="AMDAO16E05" /clone_lib="AMDOA" /note="end : T7"					
BASE COUNT	401 a      39 c      35 g      450 t      88 others					
ORIGIN						
Query Match	9.1%: Score 67.8; DB 17; Length 1013;					
Best Local Similarity	43.4%: Pred. No. 0.013;					
Matches 171:	Conservative 21; Mismatches 202; Indels 0; Gaps 0;					
Oy	26 TGTGTAAGCAATTGAATTCGTAATTTATGCTTAAACAGACCTGATGGTAGCG 85					
Db	541 WNTTATATAAATATATMTWTTATATATATTATNATATAAATTAATATWTAATATATAAAAA 600					
Oy	86 AATTCGATACAATTTGAAAAAGTTAAATTTAATAGTCCAAAAGCCTTTTTTAAA 145					
Db	601 AMATTWTATNTATTAATTAATTAATTTATTTTAAAAATTTTAAATATTTATTTMNAAT 660					
Oy	146 TTTCATTAATATGCACATAATACCAGTTTAATTAAGAAGGGTTTTATGAATAGCTTTT 205					
Db	661 AAMTTTAAATAATNAAAATWTTTATTTAATTAATMAATATATATAAATTAATTTWAAA 720					
Oy	206 AATTTTATTTTGCAATCTTTTGCTTTTCTAGTTTGCCTCAAGATGATCTTAAG 265					
Db	721 AAAAATTTANATATATAATTTATTTATTTTWTWTTTTTAAAMATANTATTTWAATTT 780					
Oy	266 CACTTTATATCGGAGCGGAGAAAAATTTTGGTTATGCAAACTAATTAAGAAAGATTC 325					
Db	781 TTTTATATATTAATTAATTTAAWATTTTTTTTTTTATTAATAATTAATAAAAAAAAA 840					
Oy	326 TCTGTACCATTTTATGAACCTTTTTTLAGGGTTCCGATAGGTCCTTTTGCCTCAAG 385					
Db	841 AAAAAAAAAAATATATAATTAATATATTAATTAATTAATAATMAATTTTWTATAAWAM 900					
Oy	386 AGATATCCTTGAGAGTCTCTTATCTTGAGATTT 419					
Db	901 AAATATTTNATTTATTTTATTTATTTATTNMAT 934					
RESULT 8						
CNSOIG8P/c	CNSOIG8P 910 bp DNA linear GSS 01-JUN-2001					
LOCUS	Anopheles gambie GSS sp6 end of clone 06el4 of NotreDamiel library					
DEFINITION	from strain PEST of Anopheles gambie (African malaria mosquito),					

```

ACCESSION      genomic survey sequence.
VERSION        AL142826
KEYWORDS       GI:7000944
SOURCE         African malaria mosquito.
ORGANISM       Anopheles gambiae
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
                Anopheles.
REFERENCE      1 (bases 1 to 910)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
                2 (bases 1 to 910)
                Roth,C.W., Brey,P.F., Ke,Z., Collins,F.H. and Weissenbach,J.
                Direct Submission
                Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
                Roux, Paris 75015, France
                This clone is from an A. gambiae BAC library provided by F.H.
                Collins and sequenced by Genoscope in collaboration with the
                Laboratory of Biochem. and Biol. Molec. of Insects, Institut
                Pasteur.
FEATURES
  source       1. 910
                /organism="Anopheles gambiae"
                /strain="PEST"
                /db_xref="taxon:7165"
                /clone="06E14"
                /clone_lib="Notredame1"
                /note="end : SP6"
BASE COUNT    376 a      83 c      91 g      294 t      66 others
ORIGIN
Query Match   8.9%: Score 66.6; DB 17; Length 910;
Best Local Similarity 43.5%: Pred. No. 0.021;
Matches 173; Conservative 27; Mismatches 197; Indels 1; Gaps 1;

QY 89 TTCGATACCAATTTGAAAAAGTTAAATTAATAGTCCAAAAGCCTTTTAAATTT 148
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 906 TTATATATATTTAGATATATATTTTATATATATATATATATATATATATATAT 847

QY 149 CATTATATGTACCATATATACAGTTTAATTAAGGGTTTATGAATGATTTTAAAT 208
    : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 846 AATATTAAMWTTATTTTBRNATATTTTATAMANANNTGTTMRTATNTNMTWAT 787

QY 209 TTTATATTTGGTAACTTTTGTGCTTTTCTAGTTTCTCAAGATGATTTCAAGAC 268
    ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 786 TTTTAAATTAATTTTATTTTNTTATTTATTTATTTATTTTATTTTAAATAT 727

QY 269 TTTTAAATCTGGAGCGGAGAAAAATTTTGGTTATGAACATAATAGAAGATTTCT 328
    ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 726 TTTTATATATATATAGATATATTAATTTTATATATATATATATATATATATAT 667

QY 329 TGTACCATTTTAAATGAACCTTTTAAAGGTTGCGGATA-GGTTCTTTGCTCAAGAG 387
    | | | | | : : : | | | | | | | | | | | | | | | | | |
Db 666 TATATATATTAAMWMTAATTTTATTTTATTTTATTTTATTTTAAATTAATAA 607

QY 388 ATATCCCTTGGAGTTCCTCTATTTCTGATTTGATGCGGTGATAGGTTAATACTTA 447
    || | | | | : | | | | | : | | | | | : | | | | | : | | | |
Db 606 TAAATTAATAAAATTTTAAATTTTAAATTAATTAATTAATTTTAAATTAATAA 547

QY 448 CAGGAGCTTATTTGACATCAAGATTTTGAATTAAT 485
    | : | | | | : | | | | | : | | | | | : | | | | | : | | | |
Db 546 AAAATATATATTTTAAATTAATTTTAAATTAATAA 509

RESULT 9
CNS00EVL/c    1101 bp    DNA    linear    GSS 04-JUN-1999
LOCUS         Drosophila melanogaster genome survey sequence 17 end of BAC:
DEFINITION    BACR29B23 of Rpci-98 library from Drosophila melanogaster (fruit

```

```

fly), genomic survey sequence.
ACCESSION      AL069706
VERSION        AI069706.1 GI:4949849
KEYWORDS       GSS.
SOURCE         Drosophila melanogaster.
ORGANISM       Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
                Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org/The BDGP Drosophila
                melanogaster BAC library was prepared by Kazutoyo Osoegawa and
                Aaron Mammoler in Pieter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of Drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's
                pl and EST libraries. A more detailed description of the library
                and how to order individual BAC clones, the entire library, or
                filters for hybridization from the BACRAC Resource Center can be
                found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
  source       1. 1101
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="BACR29B23"
                /clone_lib="RPCI-98"
                /note="end : T7"
BASE COUNT    419 a      91 c      60 g      299 t      232 others
ORIGIN
Query Match   8.7%: Score 65; DB 17; Length 1101;
Best Local Similarity 32.9%: Pred. No. 0.037;
Matches 161; Conservative 84; Mismatches 244; Indels 0; Gaps 0;

QY 15 AATAAGATATGTTTAAAGATGAATGCAATTTTAAAGTTAATCAAGAGCTCT 74
    : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 1099 WMTATATTTTATTTTATTTTATTTTATATATATATATATATATATATATATAT 1040

QY 75 ATTGTAAGCAATTTGAGTACCAATTTGAAGAAATTAATTAATTAATGATCCAAAAG 134
    | | | | | : | | | | | | | | | | | | | | | | | | | |
Db 1039 ATATATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 980

QY 135 CCTTTTAAATTCATTAATATGCTACCATTAATACCGTTTAATTAAGGGCTTTATG 194
    : : : : | : | : | : | : | : | : | : | : | : | : | : |
Db 979 TATAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 920

QY 195 AATAAGTTTAAATTTTAAATTTTGTGTTTCTAGTTTGTCTCAAGAT 254
    : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 919 TTTTATATATTTTATATATTTTAAATTAATTAATTAATTTTATATATATATAT 860

QY 255 GATCTTAAGCACTTTTAATCTGGAGCGGAGAAAAATTTTGGTTATGAACATAT 314
    | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 859 TAAATTTTATTTTATATATATATATATATATATATATATATATATATATAT 800

QY 315 AAGAAGATCTCTGTCATTTTATTAATGACCTTTTATTAAGGTTGCGGATAGTCT 374
    | : | : | : | : | : | : | : | : | : | : | : | : |
Db 799 TTTATTTATATATATATATATATATATATATATATATATATATATATATAT 740

QY 375 TTTGCTCAAGAGATATCTGTCATTTTATTTCTGATTTGATGCGGTGATATA 434
    || | | : | : | : | : | : | : | : | : | : | : | : |
Db 739 TTTATATTTTAAATTTATATATATATATATATATATATATATATATATATAT 680

QY 435 GGGTAAATACCTTACAGAGCTTATTTGACATCAAGATTTTGAATTAATGCTAAAAA 494

```

[illegible]

Df	907 T---	-BADGABGTGTCGGKCTTKGBGSKSBSBKSSBBAFRNTTTTKTNTTNTMT	854
OY	361 TC	GCGATAGAGTTCTTTTGCTCAGAGATATCCTTGACGCTCTCTATTCTTGATTTG	420
Df	853 AT	TTCCTSNYMGKGAKCKCTATRGKTRKTYCCSSNSTTGTGWMHGAACCCGTGHAAMW	794
OY	421 AT	GCGTGTGATAGGTTAATCTTAAGAGAGCTATTATTGGACANCAAGATTTTGATA	480
Df	793 A	MAANTTGHMWTWRDPTWATNCTCCCGGGGGGADPKMKGNRTDDAAAAAAA	734
OY	481 AT	AATGCTAAAAAACCTATTTAAATGTGACTTGGGGTAGAGGATGATGTTGCAGCTG	540
Df	733 A	AAAAAATGTMGAGAMMARACYCACAAGTSBCSCSSCBHNHTATATAGCGGGGTGGR	674
OY	541 TG	GTACTAGATGCGCTGTGACAAAGATTGACAGAATTTCTTTCATTTACATTTGCTAATA	600
Df	673 AA	AWATTHMAATAGAHDDMMWATTCCTACAACACTTGACTCGCTTGACCCGACACA	614
OY	601 AT	TATACACAGAA	613
Df	613 CG	CACACGACGACAA	601
RESULT 11			
CNS00DKY			
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #	BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
DEFINITION			
VERSION	AL071865		
KEYWORDS	AL071865.1 GI:4948170		
SOURCE	GSS.		
ORGANISM	Drosophila melanogaster.		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	I (bases 1 to 928)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)		
COMMENT			
	- Web : www.genoscope.cns.fr ) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES			
Source			
	Location/Qualifiers		
	1..928		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone="BACR27A24"		
	/clone.lib="RPCI-98"		
	/note="end : 77"		
BASE COUNT	262 a	70 c	84 g
ORIGIN		321 t	191 others
Query Match	8.6%; Score 64.2;	DB 17;	Length 928;
Best Local Similarity	30.4%;	Pred. No. 0.055;	
Matches	126;	Conservative	94;
	Mismatches	199;	Indels
	Gaps	0;	
5	ATTTCACAAATTAAGATATTGTTAAAGAATGAATGAAATTTGCTAATTTTATGCTTAATC	64	



Db 490 AAMATAATATGWTAAATTTTATATKWWAAAAAAMAAATWTWTWTWTWTAAAA 549  
 Qy 65 AAGAAGCTCTATGTGAAGCGAATTTGAGTACACATTTGAAAAAGTAAATTTAAATA 124  
 Db 550 AMTATTTTWTWTWTWTWTWTWTWTAAATTTAAATTTAAATTTAAATTTAAATTTATA 609  
 Qy 125 GTTCCAAAAGCCTTTTAAATTTCAATTAATGTCACCAATATGCACTTAATTAAG 184  
 Db 610 WTMTATAMAAATATWTMTATTAATATGTWTWTWTWTATTTATTTAAAMAAAAAMAA 669  
 Qy 185 GGTTTTATGATAAGTTTAAATTTTATTTTGTGTAATCTTTGTCTTTTCTAGTTT 244  
 Db 670 AAAMATATAAAAATTTTAAAMATTAAMAAAAAATTTAAATTTTWTWTWTAAATTTWA 729  
 Qy 245 TGCTCAAGATGATTTTAAGACACTTTTAATCTGGAGCGAGAAAAATTTTGATTA 304  
 Db 730 WATMAAAATATWTMMWMAATWTYDGKNNNNNNAAMMMWMMWMAAAAMMMWMMWMMW 789  
 Qy 305 TGAACCTAATAAGAAATTCCTTTGTACCAATTTTATGAACTTTTAAAGGTTGCG 364  
 Db 790 AAAAAAAMAAAAAAMAAAAAAMDDDDDDDDDDMMWAAKKKKKKKKKKKKKKKKKK 849  
 Qy 365 GATAGTCTTTGTGTCGAAGGATATCCTGTGAGGTCCTCTATTCTGCAATTT 419  
 Db 850 TKTITTKARMMTT 904

CNS020K7/C				RESULT 12
LOCUS	CNS020K7	1092 bp	DNA	linear
DEFINITION	Tetradon nigroviridis genome survey sequence T7 end of clone 22rll1 of library G from Tetradon nigroviridis, genomic survey sequence.			
ACCESSION	AL175696.1	GI:7813753		
VERSION	ALI75696			
KEYWORDS	GSS; genome survey sequence. Tetradon nigroviridis. Tetradon nigroviridis. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.			
REFERENCE	1 (bases 1 to 1092) Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence			
TITLE	Unpublished			
JOURNAL	2 (bases 1 to 1092)			
REFERENCE	Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis			
AUTHORS	Unpublished			
	3 (bases 1 to 1092)			
TITLE	Genoscope.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (12-Apr-2000)			
AUTHORS	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/tetraodon">http://www.genoscope.cns.fr/tetraodon</a> .			
TITLE				
JOURNAL				
COMMENT				

[illegible]

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	
CNS01TL7/C	CNS01TL7	1038 bp	DNA	linear	GSS 12-MAY-2000											
	Tetraodon nigroviridis genome survey sequence T7 end of clone 194p6 of library G from Tetraodon nigroviridis, genomic survey sequence.	AL166660	GI:7804398	GSS; genome survey sequence.	Tetraodon nigroviridis.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	1 (bases 1 to 1038)	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brottier,P., Quetler,F., Saurin,W. and Weissenbach,J.	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	Unpublished	2 (bases 1 to 1038)	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetler,F., Saurin,W., Bernot,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	Unpublished	3 (bases 1 to 1038)	Genoscope.
	Direct Submission	Submitted (12-Apr-2000)	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/Tetraodon">http://www.genoscope.cns.fr/Tetraodon</a> .	Location/Qualifiers												



source 1. 1038  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="194P16"  
/note="Genoscope sequence ID : COAG194DH08LP1-end : T7"  
BASE COUNT 366 a 152 c 189 g 145 t 186 others  
ORIGIN

Query Match 8.4%; Score 62.8; DB 17; Length 1038;  
Best Local Similarity 38.8%; Pred. No. 0.092;  
Matches 99; Conservative 50; Mismatches 106; Indels 0; Gaps 0;

23 TATGTAAAGAAATGGAATTCCTATTTTATGTTAATCAAGAAGCCTATGTTAA 82  
||| ||| :  
Db 813 TDTTTTTRRRRRRRRRRRRRRRRRRRATTTTTRGRRRTKRRATTTTTRATTTT 754

83 GCGAATTCGAGTACAAATTTGAAAAAGTTAAATTAATTAATGTTCCAAAAGCCTTTT 142  
:  
Db 753 RATTTTTTRRRRRATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 694

143 AATTTTCATTAATATCTACCAATACAGTTTAAATTAAGGGCTTTTATGATAAGTT 202  
||| ||| :  
Db 693 TTTTTCCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 634

203 TTTAATTTTATTTTGTGTAATCTTTTGTCTTTTCTAGTTTGCACAGATGATTTCTAA 262  
||| ||| ||| ||| : : ||| ||| ||| : : ||| ||| : : ||| : : ||| : :  
Db 633 TTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 574

263 AAGCACTTTTAACTT 277  
:  
Db 573 MMAMMTTMMTMMT 559

RESULT 14  
CNS016E2/C 1204 bp DNA linear GSS 26-JUL-1999  
LOCUS  
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC  
BACN15A12 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL106628.1 GI:5622852  
VERSION  
KEYWORDS  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydriidae; Drosophilidae; Drosophila.  
1 (bases 1 to 1204)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Genoscope.  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila genome project (ENGP) -  
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaut at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelBAC11.

FEATURES  
source 1. 1204  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACN15A12"  
/clone.lib="DrosBAC"  
/plasmid="pBelBAC11"  
/note="end : T7"  
BASE COUNT 298 a 172 c 106 g 316 t 312 others  
ORIGIN

Query Match 8.3%; Score 62.4; DB 17; Length 1204;  
Best Local Similarity 23.3%; Pred. No. 0.1;  
Matches 109; Conservative 165; Mismatches 193; Indels 1; Gaps 1;

60 AATCAAGAAGCTCTATTTGTAAGCAATTTGAGTAACAATTTGAAAAAGTTAAATTT 119  
||| ||| :  
Db 1195 MAAMATKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKK 1136

120 AATTAATTCGAAGCCTTTTAAATTTTCAATTAATATGCTACCAATTAATACCGTTTAA 179  
||| ||| :  
Db 1135 AAAAAAANKADDAADAAAAAKAKAKAKAAAAADAAKKA-AAAAKAAAKW 1077

180 AAGGGCTTTTATGATAAGTTTAAATTTTATTTTATTTGTAATCTTTTGTCTTTTCT 239  
| :  
Db 1076 KATKKDDDDKDKAAADAKKKKKAKADADAKADKDKAKADAMPDADADAAKKKKKAK 1017

240 AGTTTTCCTCAGATGATCTTAAGCACTTTTAAATCTGGAGCGAGAAAAATTTTGG 299  
:  
Db 1016 KDKKKDDADADAAAKAKKDKKKAKKDKKKAKKAKKAKKDKKKDKKKKKKKDK 957

300 GTTTATGAACATATTAAGAAGATCTCTGCTACCAATTTTATTTTAAACCTTTTAAAGG 359  
:  
Db 956 KAKKKKAAAKDAAAAAANKKKDKKDAKAAKAKKKDAKAKKKDKKDDAMADAD 897

360 TTCGGATAGGTTCTTTTGTCTCAAGAGATATCTTGAAGTCTCTTATTTCTTGATTT 419  
:  
Db 896 KAAKKKAAAKKKDKKDAADKKAADAKKKKDKAKAKADKAAKAKKDKKKAAGAARA 837

420 GATCGGTTGGTATPAGGTTAATCTTACAGACCTTATTTGACATCAAGCATTTTGGAT 479  
||| ||| :  
Db 836 AAKKGGKGGKGGAGGAGGAAAGATTTTAAATTTTAAAGTATTTTAAATTTTAAAT 777

480 AATTAATGCTAAAGAGCTGATTTTAAGTGCATTTGAGGTTTAAAGGAATG 527  
| :  
Db 776 AKWAKKTRAAAKDARKDAATTAKAARAGAAATKGGGAARAGGAAG 729

RESULT 15  
CNS035N7 576 bp DNA linear GSS 15-MAY-2000  
LOCUS  
DEFINITION Tetraodon nigroviridis genome survey sequence pUC-Orl end of clone  
214A06 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
ACCESSION AL228940.1 GI:7887933  
VERSION  
KEYWORDS GSS; genome survey sequence.  
SOURCE Tetraodon nigroviridis.  
ORGANISM Tetraodon nigroviridis.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.  
1 (bases 1 to 576)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,  
Saurin,W. and Weissenbach,J.  
Human gene number estimate provided by genome wide analysis using  
Tetraodon nigroviridis DNA sequence  
Unpublished  
2 (bases 1 to 576)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,  
Weissenbach,J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Unpublished  
3 (bases 1 to 576)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (12-APR-2000)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at

